Narrative Complexity

A Consciousness Hypothesis

Essay #3 | Dreaming

Essay #4 | Memory & Cognition

“ The grand aim of all science is to cover the
greatest number of empirical facts by logical
deduction from the smallest number of
hypotheses or axioms.”

- Albert Einstein

~

“ Of Life immense in passion, pulse, and power,
Cheerful, for freest action form’d under the laws divine,
The Modern Man I sing.”

- Walt Whitman
Leaves of Grass

What is Narrative Complexity?

Narrative Complexity is a hypothesis—a theoretical model of consciousness. The theory is built around a simple premise: that the generation & maintenance of multiple, ongoing, interconnected narratives is the primary method by which the human brain creates the flow of consciousness. This narratively-driven engine is, in essence, our mind's "operating system"—the evolutionarily-developed method by which our brain directs our choices & behavior, evaluates & analyzes the world around us, defines our place in that world, and determines how we will file, manage & respond to the seemingly overwhelming piles of data that we process every second of our existence.

Narrative Complexity explains consciousness by breaking it up into several different mechanisms that work in concert to keep our operating system up-and-running. The primary consciousness-generating mechanisms that the theory identifies & seeks to explain are:

- **Syntactic Memory: Data Storage, Association, Comparison & Recall**
- **Rules, Vocabulary & Beliefs: Narrative Building & Analysis**
- **Emotional Equations: Imprinting & Signaling**
- **Urges & Internal Dialogue: Decision-Making**

Narrative Complexity also defines & explores many of the central effects that emerge from the interaction of those mechanisms & our sensory input, such as: the Internal Dialogue Loop & the Dynamic-Core-based, multi-sensory **"Consciousness Viewfinder"** (which, together, produce the essence of our fluid conscious experience), and the **Free Will Paradox** (the essence of our sense of independence).

In addition, the theory includes a description of our levels of awareness: the **Diffuse Box of Consciousness**, which explores how these multiple narrative threads weave their way from the unaware realm of our subconscious to the state of awareness that we perceive as being.

When viewed as a whole, *Narrative Complexity* will explain how the external data that we have previously stored & are presently consuming via study & experience combines with an ongoing loop of internal data provided through our emotions, internal dialogue, physical urges, actions, speech, behavior & physical state to generate the all-encompassing, narratively-based experience of consciousness.

It's a microcosmic Theory of Everything—because, for us humans, there is no anything without consciousness.

~

Why should you care?

Because everyone loves a mystery. And *Narrative Complexity* seeks to unravel one of the great mysteries of human existence, to hunt down Philosophy's most elusive prey: consciousness itself. It explains the enigmas that define who we are:

- **How does our mind construct our experience of being?**
- **What is our brain doing "behind the scenes" when generating our consciousness?**
- **Why is our brain expending all of these resources to maintain this spectacular, fluid, ongoing state of consciousness?**

Consider this: the entity that we define as "I"—our identity, our selves—is actually our consciousness, a collection of unique phenomena primarily created & sustained within our brain. So, to answer the question of why our brain has developed consciousness is, in essence, to answer the most epic question in human history: *Why are we here?* If you want to know why you're here, you want to know more about *Narrative Complexity*.

~

Is there any proof for this theory?

Narrative Complexity is an original theory, but one whose mechanisms are well-supported by the latest neuroscience. My series of essays will demonstrate that:

- **Current neuroscience provides evidence that many of the brain mechanics & structures required to build & enact the mechanisms of Narrative Complexity not only exist, but clearly play a significant role in generating our experience of consciousness.**

- **Narrative Complexity's model of consciousness fits comfortably within the context of the most recent consciousness theories—sharing many elements with them, while also expanding or reshaping these shared mechanisms & concepts by applying a narratively-based approach.**

For those of you looking for a little name-dropping & theory-dropping to provide a better sense of which ideological umbrellas we're standing under, here's a quick list of some (but not all) of *Narrative Complexity*'s brain-brethren (others are noted within the essays):

- In terms of neuroscience, Nobel Prize-winner Gerald Edelman's work—in particular, his book (written with Guilio Tononi) *A Universe of Consciousness; How Matter Becomes Imagination* (Basic Books, 2000)—strongly supports the model of looping, interconnected neural systems proposed by *Narrative Complexity.*

- In addition, Berkeley anthropologist & neurobiologist Terrence Deacon's *The Symbolic Species* (Norton, 1999) & *Incomplete Nature: How Mind Emerged from Matter* (Norton, 2011) and Darthmouth neuroscientist Peter Ulric Tse's *The Neural Basis Of Free Will: Criterial Causation* (MIT Press, 2013) also present critical evidence that supports the neural principles underpinning my theory.

- Cognitively & emotionally speaking, the theories of psychologists Daniel Kahneman & Amos Tversky—in particular, their Nobel Prize-winning *Prospect Theory*—support aspects of *Narrative Complexity's* cognitive, narratively-based emotional mechanics. In addition, Oxford neuroscientist Edmund Rolls' *Emotions & Decision-Making Explained* (Oxford University Press, 2014) strongly supports my general approach to emotional function and our model of emotion's role in decision-making.

- Linguistically-cognitively, M.A.K. Halliday's & Christian M.I.M Matthiessen's work *Construing Experience Through Meaning: A Language-Based Approach to Cognition* (Continuum, 1999) supports my theory's view of syntax & language-based cognition.

- The model of consciousness that my theory draws most-heavily upon is Edelman's *Dynamic Core Hypothesis*. But there are other consciousness theories whose perspectives align with *Narrative Complexity*, including: Douglas Hofstadter's *Strange Loop*, Bernard Baars' *Global Workspace*, Stanislas Dehaene's *Global Neuronal Workspace*, Michael Graziano's *Attention Schema Theory*, and Ezequiel Morsella's *Passive Frame Theory*.

- In terms of the unconscious, my theory basically adopts the view presented by John Bargh & Ezequiel Morsella at Yale in their 2008 paper "The Unconscious Mind."

- And for you pure philosophy geeks: the theory label that best (although not exactly) describes *Narrative Complexity*'s approach is *Dispositionalist Higher-Order Thought Theory*.

Does this mean that you need to be familiar with the aforementioned theories in order to understand *Narrative Complexity?* **Definitely not.**

I'm not here to bombard you with talk of neurons & brain structures (although there will be *some* of that) or to spew a litany of frustratingly-obtuse, overly-used philosophical jargon like *qualia* & *dispositionalism* (there will be *none* of that).

My approach will be to use our everyday experiences, common elements of human existence, to explain & demonstrate the mechanisms that compose *Narrative Complexity*. At its heart, *Narrative Complexity* is about the experience of being human—and since you are, in fact, a human being, your own experience in that role ought to provide you with all of the background you need in order to understand the theory.

Think of it this way, not many of us could comprehend gravity if it were explained by a physicist in detailed technical or mathematical terms. And yet, based on our general knowledge of modern science, we can still understand a basic "real world" explanation of the concept of gravity: heavier objects attract smaller objects—the larger the mass of the object, the stronger the attraction. And we can use examples from our experience to provide evidence that enhances our understanding of the concept: we don't fly off of the earth, smaller objects fall toward the earth, the moon stays in earth's orbit.

Although the science behind consciousness is deeper & more detailed than ever, and consciousness' most intricate mechanics are, indeed, extraordinarily complex—by approaching the concepts from a narrative angle, we can fashion an explanation that can be understood without a degree in neuroscience.

Narrative Complexity is a journey you're already taking, this theory simply paints for you the vessel you've been aboard all your life.

~ R. Sal Reyes | April 2014

(amended & revised, September 2018 & August 2021)

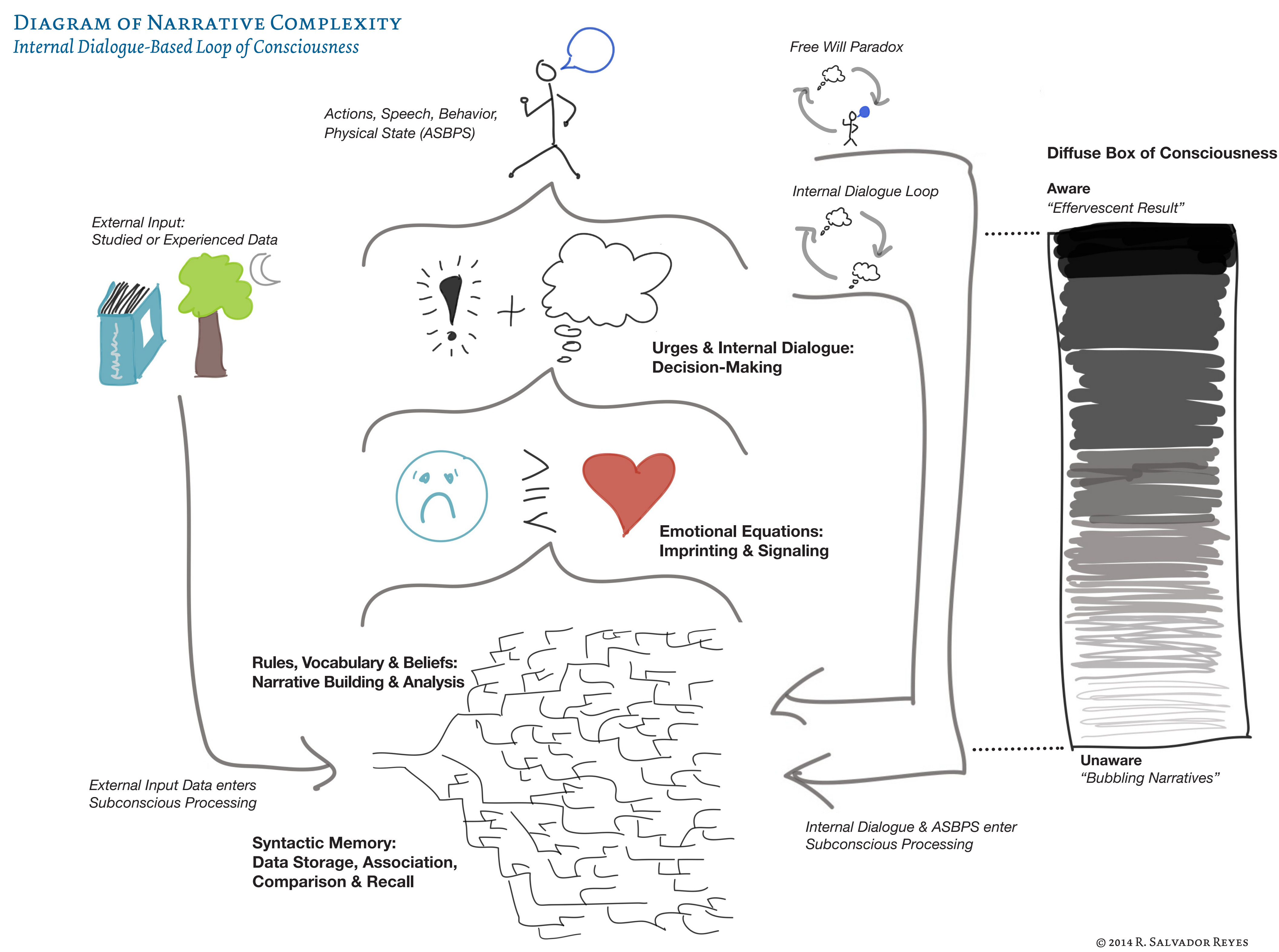


DIAGRAM OF NARRATIVE COMPLEXITY
Internal Dialogue-Based Loop of Consciousness
Actions, Speech, Behavior, Physical State (ASBPS)
Free Will Paradox
Diffuse Box of Consciousness
Internal Dialogue Loop
Aware
"Effervescent Result"
External Input: Studied or Experienced Data
Urges & Internal Dialogue: Decision-Making
Emotional Equations: Imprinting & Signaling
Rules, Vocabulary & Beliefs: Narrative Building & Analysis
Unaware
"Bubbling Narratives"
External Input Data enters Subconscious Processing
Internal Dialogue & ASBPS enter Subconscious Processing
Syntactic Memory: Data Storage, Association, Comparison & Recall
© 2014 R. SALVADOR REYES

Let's Talk About Consciousness, Baby

Good Morning: This Is Your Life

Your alarm clock begins to wail. You wake up. Your eyes take in the space: it's your bedroom, and the sun is just starting to rise outside the window. You reach for the button on the clock and read the time: *6:30 AM. I usually wake up at 7:30—why is the alarm set for 6:30? I have to pee. Oh, I woke up because I have to pee. No, the alarm woke me. Why did I set the alarm?* Your sight focuses on the jacket hanging behind the door. *My suit. The job interview: 8:30. I wanted to make coffee and breakfast first. I'm nervous, and tired: if I grab coffee & a muffin at the train, I can sleep until 7:00 and forget about the interview until then. But I should go over my notes one more time, that will relax me, plus I really have to pee: I'm getting up*. You get out of bed, feeling groggy, excited, nervous, hungry & with a strong urge to pee. Your first stop: the bathroom.

Depending on a menagerie of factors—including everything from how deeply you were asleep, to how quickly your specific brain chemistry responds to awakening, to how much alcohol you drank the night before—that series of thoughts might unfold slowly enough to hear all those words distinctly in your head, or so quickly that they barely register as sentences. But in both cases, the same basic thing has happened: your consciousness has come online. The operating system that governs every choice you will make & every emotion you will experience while awake has just booted up for the day.

And in this brief series of moments, the multiple and multifaceted interweaving narratives of that day have quickly begun to assemble, consciously and subconsciously—each building on one and other, triggering yet others, interconnecting, reassembling, submerging, dispersing & reemerging in all combinations, and all while integrating new incoming data that must be sorted, analyzed and distributed into the most relevant & useful, current or new narrative streams. Their purpose: to identify, prioritize, plan for, and seek out that day's myriad goals; and to predict the

best possible path through the day to achieve the most important and maximum number of those goals at the lowest possible "overall cost"—all while avoiding as much unnecessary risk as possible.
In this sense, each narrative is a predictive pathway toward a goal, any goal, large or small. And any moment might see any number of extremely varied goals competing for the same expenditure of resources: time, energy and/or "assets" (which is essentially the ultimate result of any choice—a decision about how we will allocate a particular available or predicted-to-be available unit of time, energy and/or assets).

When our character awoke, the urge to pee found itself in competition with the goal of wealth and prosperity that a new job might bring. The fact that our character chose to pee first instead of immediately heading to their interview notes does not indicate that they've decided urinating is a more worthy goal than wealth and prosperity. Rather, their brain was able to lay out a predicted path in which both goals could be achieved without increasing cost or sacrificing "goal value." In other words, they could pee first and still have plenty of time to do everything they needed to be optimally prepared for the interview. In fact, in a smaller calculation (one so obvious they were probably never aware that they thought about it) they might've realized that there was more likely a slightly higher cost to studying the notes first and waiting a half an hour to pee. However, if they were feeling particularly unprepared for the interview or believed it was a uniquely valuable opportunity, they might choose to "hedge their bets" and grab the notes first, then bring them to the bathroom to begin reading *while* they peed.

Urinate! Succeed! Do both! Deep down in our psyche, these are the kinds of impulses that are competing for our brain's undivided attention. Each moment of existence is a Roman Colosseum in our minds—each urge, each impulse, each desire tossed into the arena, fighting viciously to be heard, to be made part of the story, to be *expressed* out there, where the thing that thinks them acts its act in the world.

Consciousness: The Navigator

This is what our consciousness was built to do. To bring these multiple, myriad goals and all of their attendant predictive pathways into some sort of navigable focus. To provide our brain with the methodology & mechanisms needed to support humans' uniquely-evolved & dynamically-adaptive ability to interact with their environment, its creatures, and each other. To predict results and make decisions. Lots of them. Lots and lots of them, every second of every day. And to base those decisions as best as possible on data recorded in previous experiences or learned through study, and to make those decisions according to current & future circumstances & needs. *And* to access a broad, diverse array of relevant or uniquely-applicable (previously-recorded) associated data in the process of culling & sorting what specific data will be used to help generate those predictions & decisions—an associative process that is *vital* to generating *creative* or *unique* solutions to the most difficult problems that those prediction & decision mechanisms are tasked to handle. This is *Narrative Complexity*.

(That culling of the most-applicable predictive or associative data from a broad spectrum of ultimately-unrealized, but partially-recognized other data patterns reflects neurobiologist Terrence Deacon's theory about how "constraint" plays a central role in consciousness. He presents this view—and several others that *Narrative Complexity* supports—in his bold & insightful 2011 book *Incomplete Nature: How Mind Emerged From Matter.* [1])

If you think all this sounds complicated, you're right. That's why humans are presumably the first species in earth's history to possess such magnificent faculties. Whether that's truly a blessing or a curse is for the poets and philosophers to decide, but in purely evolutionary terms, it's probably the best hand that's ever been dealt on this particular blue sphere. Luckily for us, despite the extraordinary complicatedness of it all, our consciousness is also designed to keep our eye on the ball—to narrow the focus of our awareness to one or a handful of narratives that draw our conscious attention. It's a bit of a chicken and the egg argument as to whether our external attention is drawn to objects of internal conjecture, or internal conjecture arises from objects that draw our external attention; ultimately, both are happening in an ongoing fashion, and both are probably interchanging places as the "driver" of our conscious focus enough to make the process essentially simultaneous.

Nonetheless, it is this singular or nearly-singular ongoing focused narrative stream—one that combines both distinctly "spoken" & quickly "experienced" internal dialogue, terms I'll explore in more detail in later essays—that is the essence of conscious experience. (This category of consciousness, which is what our theory defines & focuses on, is often referred to as "higher-order" consciousness.) Think of this *stream of consciousness* as a narrow roadway. All narratives have attached to them an importance or urgency value. I have to pee a little or I have to pee a lot. I have an hour to get ready for my big interview or I have 20 minutes to get ready for the interview I don't care about, or vice versa.

The more urgent or important the narrative, the more "space" it takes up in the roadway of our consciousness. If your narrative is "I need to do this right now or I'll die!" your conscious roadway is pretty much at capacity. No other thoughts bubbling around in your subconscious are going to enter that narrative thoroughfare: *on-ramp closed*, *we're very busy*, *come back later*, *if we're still here*. But a few items of only medium immediate importance and that require little attention —many, like peeing, are such rote predictive scripts that they can be enacted with essentially no conscious attention allocated [2] (something we won't discuss in detail until Essay #5, *The Will Of The Free*)—a few of these lesser stories might be able to occupy the conscious narrative roadway nearly simultaneously, weaving together all their paths, and keeping all the goals "in mind" along the way. This is the navigator doing its job: circling destinations on the map and hollering directions as you go.

The Value of Deep Thoughts &
The Magic of The Internal Dialogue Loop

As we've all learned for ourselves one time or another: the worst navigators are those who keep telling you to exit here, only to change their mind after you've left the highway. Although the real-world version of this experience usually has trivial (although annoying) consequences, the narrative version can have some hidden, but very real costs. This is particularly true if your narrative goal requires some deep thinking.

A good example is the myth of multitasking (which is, frankly, an entire essay on its own, but we'll simply sideswipe it here). Multitaskers believe that our conscious roadway can simultaneously accommodate multiple narrative threads that all either require high attention or are of high interest. In reality, juggling these types of road-hogs likely requires a process that is not genuinely simultaneous or well-interwoven, but rather, is more akin to quickly sending narrative vehicles on and off the roadway to accommodate each as we switch our attention. [3]

What believers in multitasking are overlooking is the interference with narrative fluidity that occurs during this switching process, which likely hampers the brain's ability to probe the kinds of new solutions, associations & predictive models that can be accessed through a fluid *narrative loop*—primarily because this fluidity presumably allows for more extrapolations of thoughts to be processed through our unconscious.

According to our theory, in a fluid narrative loop, every thought (or thought parcel) is like an extrapolation or an echo of the *previous* thought parcel. This is one of *Narrative Complexity*'s unique & foundational hypotheses: after each sentence, phrase or idea is produced by our unconscious processing & emerges in our (prefrontal-cortex-based) conscious internal dialogue, that language-based thought parcel *re-enters* our unconscious processing (along with all the ongoing or freshly-encountered, incoming environmental & physical data).

Basically, *we hear ourselves think the thought*. And then—via linguistically-, sensorially-, and emotionally-based neural connections—that thought "pings" & compares associated, memory-stored pattern data, washes through the narrative-analyzing/building machine again, generates new or continuing emotions, aids in enacting or inhibiting any potential actions, then comes back out of the loop as the next thought on the previous one's heels. In the view of *Narrative Complexity*, this is our consciousness' primary driver, the mechanism that engages all other mechanisms that generate our consciousness: the thalamocortical *internal dialogue loop*. (Inner speech's key role in myriad aspects of experience & cognition has already been well-documented. [4])

With each loop's dive back into our unconscious processing, these ongoing extrapolations of our thoughts (essentially, the *sentences* that compose our stream of consciousness) all have a chance to ping new associations & access new patterns in our databanks for possible application and/or comparison. This is almost like a process of "thought evolution" in which increased numbers of slightly-varied iterations of an idea or thought (new sentences) allow a greater possibility of a uniquely valuable or useful response being spurred by the "pinging" of newly-associated data.

Focused attention on a series of thoughts or ideas or a narrative helps our brain to maximize these thought iterations via multiple unbroken narrative loops through our conscious expression & subconscious processing. Keep in mind that the longer a specific narrative loop goes unbroken, the more likely it is to reach a "deeper" response in terms of using multiple iterations to allow for a more complex branching of ideas. Consider that when you break that narrative loop and "return" to the thought, you are not often returning exactly to your previous location in the idea branch, but probably begin instead a few steps further back, "retracing your steps" into the idea, taking a little time to pick up speed again on the roadway and get the iterations back into that fast, fluid flow. (This tendency to begin again "a few steps back" is probably due to how the *recent & repeated recall* of that slightly older data impacts its recall likelihood—a mechanic detailed within a much deeper discussion of this entire cognitive loop in Essay #4, *You Remember You*.)

The costs of restarting each narrative might be small when viewed individually, but over time the sustained cumulative losses in the process when attempting to do something like "multitasking" can often be the difference between reaching or coming up short of the branch in the iterations of ideas where the best solution is suddenly accessible.

Language: The Creator & Ambassador of Ideas

In this looping thought-iteration process, the likely value of generating multiple, unique, cross-referenced data pings from a single thought or narrative input stream helps to explain the importance of language itself in the mechanisms of consciousness. Words are symbolic units whose core meaning is enhanced and, typically, completed by its context: the surrounding words & sentences, the real-world setting in which they are encountered, the speaker & audience, and so on. Every word represents a core expressive or descriptive value, but its full & specific meaning depends on the context of its appearance & usage.

There is an economy to this that makes sense when you think of the brain in terms of an operating system. Instead of creating multiple, large, highly-detailed units of data to represent very-specific, full versions of ideas (which would likely quickly become memory hogs & processing nightmares), it creates a core dictionary of malleable terms, and uses a system that allows these terms to build a full idea's specific details through a complexity that emerges via the interaction of the core terms.

Thus, words are just malleable enough to be highly-varied & dynamically-applied in their usage (therefore, more frequently useful), and yet just solid enough in their core meaning to allow for a mostly-predictable, un-confusing, specific result in that same dynamic usage. Therefore, instead of having one word that only & specifically means "I see a red snake by the river this morning," and another word that only & specifically means "I see a green snake by the river this morning," we have eleven less specific words that can be combined to say either, or a plethora of other very specific things.

The human brain's ability to build thoughts & ideas with interchangeable, highly-configurable units capable of multiple associations and usages became possible through the development of our neurons and associated brain structures. Deacon's theory of language evolution (presented in his brilliant & provocative 1999 book, *The Symbolic Species* [5]) suggests that the evolution of the primary brain structures & capacities required for language actually occurred over an extended period of time in mammals before humans emerged (due to non-language-based evolutionary forces).

And it appears that as these mammalian brains evolved, their neurons essentially developed those robustly modular, programmable (& *re*-programmable) abilities that allowed the complex creation &

analysis of the cognitively-generated predictions & choices that would eventually empower & be required to manage language. These abilities were partly acquired through neurons' ever-increasing capacity for more & different types of connections between each other. (As we'll explore next essay, our *emotional* cognitive systems also benefitted from advancing neuronal capacities—e.g., our unique & powerful *spindle neurons*, which only appear in humans, great apes, elephants & cetaceans, and are present in brain regions like the anterior cingulate & fronto-insular cortexes—areas that appear to be heavily involved in emotional analysis.)

The existence of these kinds of neural structures & their looping, highly-malleable, powerfully-associative capacities is supported by the work of Nobel Prize-winner & neuroscientist Gerald Edelman (& his frequent partner Guilio Tononi)[6]. In addition, those "*re*-programmable" neurons are central to pioneering neuroscientist Peter Ulric Tse's theory of "Criterial Causation" (he dubs it *synaptic resetting*). He explains this mechanic in his groundbreaking 2013 book, *The Neural Basis of Free Will: Criterial Causation*[7], which presents a powerful case for the neural properties & mechanisms required by *Narrative Complexity*'s systems of cognition (the focus of *Essay #4*).

Additionally, in the view of our theory, all of those neural mechanisms identified & defined by Deacon, Edelman, Tononi & Tse are exactly the kinds of brain systems necessary to support & manage the model of language-based cognition proposed in M.A.K. Halliday's & Christian M.I.M. Matthiessen's seminal 1999 book, *Construing Experience Through Meaning*[8]. Their deep, complex & pioneering theories of language & grammar strongly support our own hypothesized syntactic systems & the consciousness-sustaining language-based cognition process that we are proposing here (& will discuss much more deeply in *Essay #4*).

Returning to that prehistoric path of mammalian brain evolution—by the time primates arrive on the scene, they are capable of using their evolved, modular, complex systems of cognition & behavior to develop sophisticated & dynamic responses to many unique & complicated problems. But they cannot manage these responses with that extraordinarily powerful & symbolic tool: human language. In essence, they have no real words & thus no internally malleable way to represent & symbolically cross-associate all those modularly-constructed, wordless-yet-dynamic (& rudimentarily "creative") thalamocortically-based behavioral responses. It is not until hominins developed their highly-unique & sophisticated control of vocal faculties (again, due to primarily non-language-related evolutionary forces) that they were able to begin developing & nurturing complex human language (something else that Deacon details in *The Symbolic Species*).

Thus, it's likely that language took hold of those already sophisticated cognitive systems incrementally—with language itself & hominins' slowly-refining/repurposing brain structures/mechanics each helping to push our ancestral minds along the path toward modern human linguistics (a process that Deacon describes as "co-evolution").
Going back to those pre-mammal minds, think of it this way: sharks, amphibians, reptiles, and other simple-brained creatures of their ilk are all essentially what we would consider robots. By this I mean that they basically have fixed responses to very specific data input, almost all of which has been pre-programmed. If external input satisfies some, but not all of the specific "data-point" requirements for a pre-programmed fixed response, the response will not be triggered.

This leads to highly-controlled, highly-predictable (thus, more reliable) behavior, but it does not allow the creature to adapt very well to its environment. Essentially, these robotic brains have a severely limited ability to learn & distinguish the similarities & differences between like-but-not-identical patterns, and therefore possess a limited ability to dynamically combine any component parts of previously learned data for use in new situations.

In the view of our theory, this is, at its core, a result of the creature's neural limitations. Based on the highly-specific, pre-programmed, robotic & non-dynamic (essentially, entirely *reflexive*) nature of their behavioral responses, those "early" creature brains do not appear to have the types of neural structures required to respond to & record experiences (in essence, *ideas*) in a complexly modular (independently-associative component-based) & creative fashion. Thus, these creatures cannot compare and connect the component parts of a data pattern—because most integrated, multi-modal experiential data patterns in early brains likely have essentially *no* independently-associative component parts. (*No modular experiential data structures.*)

Their operating system is still using that reliable, but clunky and old-fashioned method: *one word that only & specifically means "I see a red snake by the river this morning."* In fact, for much of the creature kingdom the operating system is even more rudimentary than that. Their method is more like: *one word that only & specifically means "I see red; now run."* Obviously, these creatures don't literally have "words." But they do have neural structures that correspond to experiential-data patterns and are used to help determine the creature's responses—which is ultimately what human words & language are.

Beginning in amniote (i.e., reptilian) brains, it appears that rudimentary, non-modular, but *remembered* or *learned* experiential-data patterns—high-pain experiences that became the earliest forms of simple memories—were handled by the *amygdala*. (This system still, in fact, exists in humans, which we'll discuss in Essay #4.)

As we move up the evolutionary brain ladder, growing sophistication within (& more sophisticated relationships *between*) areas like the cerebral cortex, amygdala, basal ganglia, thalamus, hippocampus and cerebellum allow for more robust (*cortex*-based) memory & learning mechanisms to be added to the operating system in creatures like birds and mammals. (Although birds' neural systems diverge in some distinct ways from mammals, their advanced methods of data-handling mimic many mammalian mechanisms. And, in fact, recent research has shown that highly-intelligent birds like crows demonstrate creative, communicative & behaviorally-sophisticated cognitive capacities that are comparable to advanced mammals.)

But many of the earliest mammalian learners were still limited by their inability to construct truly complex, modular, multi-association experiential-data patterns within these cortex-based memories. Thus —although cortex-based memory & learning mechanisms in early mammals (like mice) are more complexly, broadly, frequently & usefully applied than those simple reptilian amygdala-based mechanisms—early mammalian learning is still mostly limited to basic pain/pleasure encoded responses to either a large, very specific non-modular data pattern (*"When I see a red snake by the river in the morning, run"*) or a single data point (*"When I see red, run"*). This means that the next time either of those little-minded fellows comes across a dusty-green rattler in the desert for the first time, they're probably screwed. (As we'll discuss later, early modularity in these record/response neural mechanisms probably began with mammals like early canines or even humble guinea pigs—although those systems are far less complex than the systems that emerged in primates.)

Humans, on the other hand, have an operating system that can say in its modular, multi-word way *"I saw a red snake by the river this morning and he bit me,"* and then later say *"I see a green snake in the sand."* Here the common modular element "snake" connects the two ideas and the data from the potentially life-threatening earlier experience is pinged & cross-referenced, spurring a new narrative response that leads the human away from the danger. Believe it: *words save lives*.

Or, to view it in less dramatic terms—like saving memory space, and allowing for more malleable, dynamic, interchangeable units of idea construction—the benefits of symbolic, adaptively-configurable words over highly-detailed, idea-specific words are fairly obvious. But the hidden value of this type of symbolic language, and its special use in our consciousness' internal dialogue mechanism goes back to our discussion of iterations of thoughts.

Because each word has multiple uses in multiple settings, every time it enters our subconscious processing via internal dialogue, there is a greater possibility that in this new context the word's multiplicity of connections will help generate one of those "uniquely-useful" pings of a now suddenly-associated, formerly-unlinked idea or piece of pattern data (a cross-matching capacity that is, by comparison, *severely limited* in even highly-advanced & cognitively-creative but *non-complex-language* mammals like apes).

In addition, thanks to that malleable power of language, this multiplicity in now-comparable connections can aid in the creation/discernment of a broader, more useful *symbolic pattern*. Another way to look at it, word *malleability* (usefulness in multiple contexts) allows & enhances both more-direct "*snake-to-snake*" connections between different ideas/experiences, *and* less-direct, more-symbolic (and more-broadly-useful) "*snake-to-guy trying to steal my girlfriend*" connections between different ideas/experiences.

Thus, it is because of the malleability of words & their symbolic content that they are able to bridge the gap between larger concepts that might otherwise remain unconnected if compared as wholly-constructed, complex idea patterns. When these complex patterns are linked by a singular or a subset of common modular component part(s), the connection between them and the possibility of cross-application & larger symbol generation/discernment becomes possible. *It's the power of metaphor*. This kind of useful pollination between incidentally-relatable but seemingly-unlike larger ideas is the root of human creativity, the very essence of the problem-solving virtuosity that has propelled humanity to such dizzying heights.

Internal Dialogue: This Is Who You Are

Although it happened so quietly that you probably didn't even notice, we just answered that celestial question-of-questions: *why are we here?* Which is really the question: *why this internal dialogue shtick?* Upon first glance, it seems that human beings could function in a highly-complex learn & adapt fashion without experiencing the manifestation of an observational & conversational internal dialogue. This dialogue-less creature could use the same modular data structures to record & encode new data, then connect & compare it, etc., generating a seemingly similar range of behavioral & action responses—all without that experience being reduced to one or a few internally "spoken" key narrative streams. This would seem to be a creature very similar to a human in all outward ways—

except that it probably wouldn't talk, which quickly reveals one of the creature's flaws, and one of the basic benefits of words. Social behavior, cooperation, negotiation—some of the most crucial interactive tools contributing to human advancement seem nearly impossible without language.

But even in those social arenas, there are still less-costly evolutionary developments that could have supplanted the role of words in aiding our progress. It's not hard to imagine that rudimentary sign language (something much more akin to pantomime than modern word-based signing) and other forms of non-word-based communications could have been powerful drivers in the area of social interaction and allowed plenty of human advancement before there was any real evolutionary pressure to make the complicated & spectacular leap to an internal dialogue capable of sustaining the experience of consciousness.

And by supplementing those pantomimes with rudimentary, vocalized, word-based language (which is likely what occurred) it seems that early hominins could have developed an even more useful system of communication that *still* doesn't require complex self-sustaining internal dialogue to access many of those early cooperative & social benefits. Thus, if this interactive social aspect of rudimentary language was its primary (or only) evolutionary advantage, there would not seem to be any powerful push for it to evolve into that spectacular system of complex language & internal dialogue.

There is, in fact, evidence of a modern group of humans who have built a rudimentary "language" from such non-word-based pantomiming: a group of deaf individuals in Mexico who never learned sign language and who communicated via basic, communally-shared & -developed pantomimes (depicted in Susan Schaller's 1995 book *A Man Without Words* [9]). Interestingly, even after one of these individuals eventually learned word-based signing, he basically couldn't provide any kind of language-based depiction of what it was like to live without words; he referred to it as a "dark" time, a confused former state that he had no desire to describe. In essence, according to our theory, without words he was unable to generate a fully perceivable & recount-able conscious experience—resulting in that "dark" time of an amorphous, confused, wordless & thus, narrative-less existence.

Obviously, hominins did not remain in such an internally wordless, self-dialogue-less state—there was clearly strong evolutionary pressure for our brains to make that spectacular leap forward. Translated: there was a very rewarding advantage created by inching generation-by-generation, mutation-by-mutation, toward a brain that talks to itself using words, toward an internal narrator. And it is that internal narrator—the one that says "I am here," even in total sensory deprivation, as long as the brain is conscious, or at least semi-aware—it is that internal dialogue that truly defines us as *us*, as the thing that is our self-aware "being." [10] We know innately: I am here if I can say to *myself* that I am here. That mere snippet of internal dialogue is the essence of being: *I am here.*

This is the key to understanding the *uniqueness* of human consciousness: once our systems of dynamic language production have been learned (our toddler years) human consciousness is—at its most fundamental & unadorned core—essentially *entirely unrelated* to external sensory input. As long as a human has language, even in the *total absence* of external sensory stimuli, internal dialogue can continue to *self-generate* dynamic cognitive responses (creative self-sustaining thoughts) via its perpetually-looping nature.

I briefly experienced just such a sensory-deprived, but linguistically-conscious & coherent state prior to fully emerging from a seizure-induced unconscious episode that occurred in my late 30s. The thoughts I had in the those minimally-conscious moments—which presumably occurred while lying in the emergency room bed, unable to move or feel or hear or open my eyes—mostly focused on wondering what kind of dire predicament I'd gotten myself into, and whether or not I was, in fact, dying. In this state, I still retained a good understanding of who I was and the general facts about myself, but I was lost in time—unable to remember where I was in the story. (This *lostness* is something that would actually continue for several weeks after I awoke—a result of the temporary amnesia caused by the seizure, which I'll discuss more in Essay #4.)

In the view of our theory, this kind of sensory-deprived but coherently-conscious experience *is not possible* in any other (or *non-linguistic*) animals—even other advanced mammals, whose dynamic "cognitive" responses *require* sensory data to be constructed, because without internal dialogue *there is no other source* of useable incoming data. (The rare exceptions are possibly creatures like highly-advanced & rudimentarily-language-capable cetaceans—i.e, dolphins, which makes humans' oft-horrific treatment of them even more disturbing to contemplate.)

In the absence of sensory stimuli, a pre-language mammal brain might *attempt* some kind of cognitive behavior generation by essentially randomly associating their "darkness" to stored data & engaging cognitive processes in that way. But because their lack of internal dialogue makes them incapable of "narratively-contextual" cognitive rule application, these brains require that fundamental spatial/physical context (absent in this scenario) to effectively choose which cognitive rules to apply to that random data—which means any attempts at cognition would essentially result in useless behavioral nonsense. Basically, these creatures would be reduced to a waking version of their *dream state. (*Dreams are the subject of Essay #3, *The Night Shift*, which hypothesizes that language-less mammals' dreams are likely a nonsensical, narrative-less, non-contextual internal replaying of incidentally-associated experiential data: *a flash of chasing followed by a flash of eating, and so on.*) This waking version a non-contextual *dream state* in a language-less mammal would be entirely unlike the robust & reasonable "awake-but-sense-deprived" *internal dialogue* that humans can experience—even in that *total absence of external stimuli*. (We will explore a deeper comparison of human & pre-language mammalian cognition is Essay #4.)

This means that all of that rich, detailed fully-integrated sensory data that we experience via our "consciousness viewfinder" of awareness (yet another concept we'll explore in Essay #4) and which *seems* central to human consciousness is really just a pre-packaged (& extraordinarily-useful-to-the-point-of-near-necessity) system of external data processing that comes built-in to mammalian brains because it was once the *only* data source for cognition. In humans, however, that system of external data processing is not actually necessary to run our system of language-based internally-self-sustaining & dynamic conscious cognitive processes. (How that internal dialogue loop manages to be effectively *self-generating* without some essentially metaphysical *self-entity* directing the focus & scope of that dialogue is explained in our discussion of "narratively-contextual rule application" in that frequently aforementioned Essay #4.)

The reason we strongly, desperately *prefer* to run our language-based consciousness system along with this rich sensory input system is that it allows our consciousness to actually do *useful stuff* with its cognitive powers—like responding to that sensory environment to satisfy our needs and correlate sensory data to internal dialogue that is simultaneously being integrated within that *in-the-moment* experiential arena that's anchored by our prefrontal cortex.

This internal dialogue capacity is so powerful & central to humans' conscious "being" that even if we have completely forgotten who and where we are, we will and *can* still tell ourselves that essential fact: *I am here*. Anything less is viewed as unconsciousness or consciousness without "being" (or without any form of "being" that would be recognizable to us in a line-up). Dualism's silliness might've been beyond Descartes' grasp (and really, who could blame him—it sure *feels* like there's some kind of *floaty thing* inside this other more obviously visible & awkward one) but he really nailed it when he conjectured: *I think*, *therefore I am*. You just can't argue with it.

In fact, thanks to our brain's (very useful) obsession with cognitively *mapping* most of our internal data to some part of our body, internal dialogue is likely why we sense that "floaty thing" thinking *inside our heads*. (In other words, we don't sense that this floating voice is inhabiting our hand or our leg.) As we'll describe in the next essay, our brain likely maps our emotions to different body parts—which is why we sense that *feely thing* inside us. Similarly, we sense that *floaty thing* in our mind because our brain likely maps internal dialogue-based auditory data to our head. (And the internal dialogue also helps us to cognitively contemplate both those *feely & floaty* parts of ourselves.) Essentially, the brain is trained to map almost *all* (consciously-experienced/modeled) internal data to *some* part of the body (otherwise, generally speaking, that data isn't much use).

All of which tells us *how we know* (or *sense*) that we're here, but the question I promised was: *why* are we here? Why *did* those early humans end up with little voices in their heads instead of remaining modularly-cognitive, pantomiming & rudimentarily-speaking, but internally dialogue-less zombies? The short answer: problem-solving virtuosity. It's all about the loop, baby! (A sentiment that's shared by cognitive scientist Douglas Hofstadter, who pioneered the idea that a "strange loop" is the centerpiece of consciousness, and whose work I greatly admire. [11, 12]) As we'll explain in Essay #4, the beauty of a sentence or a thought is that it's essentially a dynamically-created symbolic equation. It's a type of mathematical hypothesis, either an *observational* or *causal* pattern of sorts—one that our brain tends to view as a problem to be solved in some way, or maybe more accurately, as a proposition to our sub-conscious processing: *whaddya get from this?*

(Although thoughts are ultimately presented in that linear fashion typically associated with our "left-brain," the mechanisms that lead to this linear product occur in a primarily parallel pattern-processing fashion. In other words, the brain does not function like an algorithmic & linear "computer"—in *very*, *very* simplistic terms, we might think of it instead as a powerfully-associative, heuristically-oriented pattern-matching & -processing machine.)

When a thought from our internal dialogue is reabsorbed into the subconscious, this "equation" or *observational/causal pattern* and its data are basically being submitted for a quick-but-thorough, cavity-probing Google search of the brain's vast memory-based data banks. And as we noted earlier, it's the metaphorical, transitive abilities of symbolic language that unlock the cross-referencing, cross-application, problem-solving power within these data banks.

There are a couple of killer-app-like advantages to using this system of generating & reabsorbing a narrow stream of word-based, narrative thoughts that pertain to your area of attention. One, by using a method that sends only the highest priority or most attention-requiring narrative(s) into the internal dialogue loop, the brain is sorting and guiding the momentarily most important, relevant or useful current data into the part of the system that has the necessary & devoted resources ready for high-powered Googling & cross-checking. Two, before that Googling & cross-checking, each sentence or cycle of internal dialogue is reconfiguring the complex, high-priority data of the moment into the more-efficient symbolic terms crucial to the useful cross-application of pattern data.

If our brain didn't reabsorb this stream of word-based thoughts, that circularly-looping data pathway would look more like a *U-turn arrow*: vacuuming in environmental & physical sensory data at one end, processing it linguistically, then launching the resultant word-based narrative parcels out the other end & into the world via speech, but never allowing the brain to make *internal* use of all the syntactic & vocabulary-based data contained within those parcels. Such a "*speak-but-don't-think*" (*aka*, zombie-like) system would, thus, be bereft of all those killer-app-like symbolic & associative advantages provided by our system of *looping* self-heard internal dialogue.

Why are we here? Because a brain that talks to itself is likely to be much, much better at coming up with unique solutions to our most pressing and/or most difficult problems. And those crazy-sounding, echoing-in-your-cranium musings also help your brain to focus its problem-solving mechanisms on the most crucial or immediate matters in our purview, thus ensuring that the brain's most useful processing resources are being devoted to analyzing the most important data. Of course, "crucial," "immediate," and "important" are *very* relative terms, depending on the particular cranium that's doing the musing. (This is a matter we'll discuss in great detail in Essay #2, *Monkey Feel, Monkey Do*, which covers our human *emotions*; I believe our emotions' evolved, in-born, complex gain/loss & prediction judgement systems provide the value- & validity-based data-encoding & behavioral-signaling required by those cognition-producing neural mechanisms proposed by Edelman, Tononi & Tse.)

Although the very earliest usage of language among hominins was, indeed, likely driven by social, verbalized person-to-person exchanges—because of the way self-produced speech is integrated into our systems of perception, it would not have taken long for the simple repression of actual vocalization to produce the first sparks of internal dialogue. Over time, this capacity to run our complexly-useful &

creative language process in an ongoing fashion quickly helped make the many uses of internal dialogue (& its recruitment of neural resources) grow exponentially—eventually becoming the dominant component of human consciousness.

Ultimately, you're here because without you, your brain might never realize that a *bucket* isn't just "a cylindrical, topless object that can be filled with and dispense water," but rather, that a *bucket* is "a device for carrying stuff." A thought which—many eons after buckets were actually invented—might've helped give somebody an idea when they were building a system for programming computers and wanted to make some of this mass-less stuff easier to handle in their little system, and they were thinking "y'know, like to *carry the stuff* around...wait, like in a *bucket*, I'll make virtual *buckets*." That's why you're here—to create *buckets* from *buckets*. Sure, it doesn't sound very romantic, but it did make evolving toward our conscious existence seem like a good idea for our species, so it has to get some props for that.

All Narrative, No Complexity Makes Jack a Dull Boy

Happily, despite the underwhelmingly pragmatic foundations for the development of consciousness, romance is never far from the human mind. And the same evolution of neurons & neural structures that allowed for symbolic language and modular data systems also mirrored the evolution of our more romantic consciousness-generating faculties: our capacity for sophisticated memories, complicated belief systems, and complex emotions. All of which we'll explore in delicious, passionate detail in later essays, as well as some of the more swoon-worthy side-effects of our oh-so-functionary, consciousness-inducing internal dialogue, and a few other secrets that will have to be deviously kept for now. (We'll also explore the ways in which other vertebrates' *language-less* conscious experience, emotions and cognition are *very similar* to our own.)

Until then, a final word about the final word in Narrative Complexity. The *complexity* is all in the neurons & the language they enable. It's in their combined, magnificently-evolved ability to freely connect, associate, compare, extrapolate, reduce, measure, encode, discard, assemble and disassemble all the data taken in and subsumed by the human brain during an entire lifetime. Without our complex neurons & language, we would be those *thoughtless* zombies. Things that didn't think they had that *floaty thing* inside this other visible thing. We'd be things that didn't *think* at all. And of course, as we all know...*one more time, with feeling:* I think, therefore I am.

###

Endnotes:

p.8:
1. Deacon, Terrence. *Incomplete Nature: How Mind Emerged from Matter*. Norton, 2011.

p.9:
2. Bargh, John A., and Ezequiel Morsella. "The unconscious mind." *Perspectives on psychological science* 3.1 (2008): 73-79.

3. Dux, P. E., *et al*. (2006). Isolation of a central bottleneck of information processing with time-resolved fMRI. *Neuron*, 52: 1109-1120.

p.10:
4. Perrone-Bertolotti, Marcela, et al. "What is that little voice inside my head? Inner speech phenomenology, its role in cognitive performance, and its relation to self-monitoring." *Behavioural brain research* 261 (2014): 220-239.

p.11:
5. Deacon, Terrence. *The Symbolic Species*. Norton, 1999.

p.12:
6. Edelman, Gerald M., and Giulio Tononi. *A universe of consciousness: How matter becomes imagination*. Basic Books, 2000.

7. Tse, Peter Ulric. *The Neural Basis of Free Will: Criterial Causation*. MIT Press, 2013.

8. Halliday, M.A.K., and Matthiessen, Christian M.I.M. *Construing Experience Through Meaning: A Language-Based Approach to Cognition*. Continuum, 1999.

p.16:
9. Schaller, Susan. *A Man Without Words*. University of California Press, 1995.

p.17:
10. Morin, Alain. "Possible links between self-awareness and inner speech theoretical background, underlying mechanisms, and empirical evidence."*Journal of Consciousness Studies* 12.4-5 (2005): 115-134.

p.19:
11. Hofstadter, Douglas R. *Gödel, Escher, Bach.* New York: Vintage Books, 1980.

12. Hofstadter, Douglas R. *I Am a Strange Loop*. Basic books, 2007.

Monkey Feel, Monkey Do (And Vice Versa)

Star Trek's Big Lie

Emotion. The muse of the volatile and irrational. The enemy of reason. The Yin to Logic's Yang. *Or so our culture says.* To wit, this dichotomy is a primary theme of possibly the greatest (& most ponderous) cultural artifact of our era, the *Shakespeare* of the late 20th century: Gene Roddenberry's *Star Trek* (in all its incarnations). Vulcans, androids, cyborgs, holograms—each is a science-fictionalized projection of a core modern human belief: that submitting fully to logic is synonymous with abandoning emotion, and vice versa. I'm here to tell you: *they've got it all wrong.*

In many ways, emotion is *pure* logic. Or, more accurately, it's pure logic cut with a dose of gambling. But to understand why that's true, we need to begin with the original purpose of emotions. In the simplest terms, mammalian brains first used emotions to tag basic pattern data (essentially, things & events) as helpful or harmful. Over time, evolving neural structures have allowed our feelings to reflect more complex judgements, but at their core they're all still designed to trigger the same binary response: inhibit or encourage an action/behavior. *Ouch! That red glowing stuff is hot. Mmmm! This stuff I'm eating is yummy.*

Pain. Pleasure. The ancestral root & ultimate result of all feelings. Forget the false *Star Trek* dichotomy of logic & emotion—whose purposes are nearly identical—the real Yin & Yang of our minds is Pain & Pleasure. Without them, the human brain would almost be *incapable* of exercising logic. Think of it this way: when we say we want to make a decision *logically*, we're essentially saying that we want to make that decision strictly by weighing our choices' *most-believable cost/benefit ratios*. More conversationally: logic is all about reducing decisions to *the cold, hard facts of the matter.* But in our brain's predictive and decision-making equations (those interweaving narratives in our mind) our emotions *are* the cold, hard facts—the fixed values that our brain uses to calculate each choice's most-believable (*highest validity*) cost/benefit (*pain/pleasure*) ratio.

The Logic of Emotions

Imagine that we accidentally dropped that aforementioned *yummy stuff* into that *hot red glowing stuff*—our brain has a choice to make: *do we tell this clumsy idiot to reach into the fire for his last piece of newly-discovered yummy or do we make him cry over its loss?* To make this choice the brain likely (via internal dialogue) quickly tells itself at least two stories (unless it has a closely-related & well-remembered previous experience to call upon for a more reflexive response). Each story is one of those predictive, decision-making (and *emotional*) equations that our mind is perpetually calculating. The narratives might go something like this (although in any specific case, obviously, the actual "heard" syntax might be far more simple or detailed):

1 - *Idiot reaches into fire, burns hand briefly but harmlessly (small value loss), retrieves yummy & consumes (medium value gain), and feels pleasure. (Narrative pattern is tagged with this pleasure—whose future purpose is to encourage reaching into small fires for medium value assets.)*

2 - *Idiot watches yummy burn (medium value loss) and cries, feels pain. (Pattern tagged with pain—unlike story #1, this event is probably not categorized as its own narrative. Instead, it's seen as the final plot twist in the story "being careless while eating something excitedly over a fire" and thus, this pain's future purpose is to inhibit such situational carelessness. Additionally, I believe that the "lightly experienced" emotion generated simply by running this predictive scenario in your mind after dropping the yummy encodes the actual memory data with enough pain to mildly discourage future situational carelessness, even if you choose to retrieve the yummy and never experience the pain of actual loss).*

After quickly comparing these two predictive narratives, the brain is most likely to lead the idiot to retrieve the yummy and achieve pleasure. In order to foresee that pleasure, the brain needed to calculate the net result of the predicted value loss & predicted value gain. These *values* are partly derived by the intensity & type of *emotion* (pain or pleasure) experienced when the data was first tagged *Ouch!* or *Mmmm!* Thus, the "emotional equation" of story #1 is something like: burn pain (-1 value, partly derived from *Ouch!*) + yummy consumption pleasure (+3 value, partly derived from *Mmmm!*) = net pleasure (+2 value).

Note, however, that I said these values are only *partly* derived by the strength of the original pleasure/pain tag. That's because this value is actually likely the result of a "sub-calculation" that combines *three* basic judgements of a narrative event or element: importance, relevance, and novelty. *(In Essay #5, we'll discuss more about how the brain makes these "Narrative Prioritizor Test" judgements & how they impact decision-making.)* In story #1 the *importance* of consuming the yummy is determined by that original pain/pleasure tag (really tasty & satisfying *Mmmm!* signals greater gain and equals higher importance).

The *relevance* is determined by the fact that it is the idiot's yummy, therefore highly relevant. If he intended to share the yummy with, say, a random wanderer who just dropped by, the gain is essentially half as relevant (but if the sharer is, instead, part of his family, the gain might still retain high relevance).

The *novelty* is determined by several things here: it was the idiot's last piece, it was the first time he'd ever found this yummy, and he does not believe these specific yummies are in local abundance. This all gives it high novelty, further increasing the yummy consumption's total pleasure value.

This sub-calculation not only determines the full value of that specific narrative event (idiot consumes yummy=+3 value) but ultimately helps determine the value/intensity of the net pleasure generated (both predicted and actual) as a result of the full narrative (idiot reaches into fire, burns hand, retrieves yummy & consumes=+2 pleasure). And the brilliant Daniel Kahneman's & Amos Tversky's Nobel Prize-winning *Prospect Theory* has shown that our brain is calculating these exact kinds of complex, predictive, *contextually-defined* gain & loss computations (much more complex than this one) when making those decisions that our consciousness governs. [1, 2]

Although Kahneman was awarded the Nobel in Economics, Prospect Theory's insights actually apply to how humans judge risks & rewards in all kinds of decisions, not just financial ones. And in the view of *our* theory, it's clear that the human brain's emotionally-based value gain/loss judgement mechanisms don't distinguish between "monetary" gains/losses and gains/losses of all other kinds of resources (time, effort, non-monetary assets, social capital, personal support, affection, etc. *ad infinatum*).

Our brain's emotional & decision-making calculations ultimately *don't care* what the actual *substance* of the gain or loss is. Either it did/could *help us* or it did/could *harm us*— and the contextually-determined degree to which we judge it did/could *help/harm* is the data that our brain uses to judge how much value (of any kind) has been (or will be) gained or lost. (At the root of these value judgements is that just-mentioned *importance/relevance/novelty* "Narrative Prioritizor Test.") Thus, when Prospect Theory demonstrates how humans "feel" about & calculate those risk/reward decisions based on contextual (narrative) gain/loss predictions, the theory is demonstrating how humans calculate *all* feelings & decisions about contextual gain/loss predictions & events: aka, *emotions*.

In Narrative Complexity's model, this gain- & loss-based "emotional analysis" of narratives occurs near the tail end of our cognitive loop. We'll discuss cognition & emotions' role in decision-making in Essays 4 & 5, but in our model, this (highly-diversified) emotional analysis occurs just after a language-based narrative parcel has been neurally constructed (& *prior* to that parcel entering our conscious awareness). The results of this analysis (which involves myriad brain areas, including the anterior cingulate cortex, orbitofrontal cortex, insula & amygdala) are routed to our decision-making *Dorsolateral* PFC (to help determine the activation/inhibition of actions) and to the appropriate emotional-response areas (e.g., the hypothalamus), which aid in producing emotionally-based bodily responses and "feeling-producing" neurotransmitter/hormone output. (The general principles of our model of emotional mechanics & emotion's role in decision-making are strongly supported by Oxford neuroscientist Edmund Rolls' recent groundbreaking work, *Emotions and Decision-Making Explained* [3].)

In addition, according to our theory, these narratively-produced emotions are ultimately routed to our somatosensory cortex, which helps us to actually *perceive* our emotions. The somatosensory cortex is involved in processing tactile sensations (including physical pain) and mapping those sensations to specific locations in our body. For example, the insula receives tactile information such as physical pain (from the nervous system via the thalamus) and likely uses its connections to the secondary somatosensory cortex to send those insula-processed pain (or pleasure) judgements to that somatosensory system for mapping to a specific part of our body (the insula also sends instructions to the hypothalamus to help produce those neurotransmitter/hormone-fueled responses). Similarly, we hypothesize that the insula also receives narratively-produced syntactic & semantic data (which contains the content required for those "emotional equations") and routes its emotional analysis of that data to our somatosensory cortex, allowing us to physically feel & perceive the emotion.

Of course, emotional pain & pleasure don't directly correlate to specific body parts. Nonetheless, because this somatosensory route is essentially the only way that our brain can physically map & perceive "feeling" an emotion, those narratively-produced feelings still seem to be experienced in (sometimes vaguely-defined) areas of our body. And I believe that the bodily area in which we feel an emotion generally correlates to the part of the body associated with the primitive, root "proto-emotion" from which that emotion evolved.

We'll discuss these proto-emotions in great detail near the end of this essay, but the simplest example is the proto-emotion that we hypothesize is the root of all basic pain/pleasure: *hunger/satiation*. Hunger/satiation is obviously a sensation felt in (& cognitively mapped to) our stomach—thus its evolutionary-descendent, emotional pain & pleasure, is often also felt in (& cognitively mapped to) our stomach. In the view of our theory, this mechanic is a key element of how we experience the many-varied emotional states that can be produced by our consciousness-sustaining internal narratives (additionally, I believe that we can *learn*—via experience & study—to associate different body parts with different emotions, which can also impact how we cognitively map & perceive these feelings).

Ultimately, all of this means that emotions are *not* some separate neural mechanism that is *competing* with our more "rational" cognitive processes (that *competition* is actually provided by our more primal *urges*, aka, those aforementioned *proto-emotions* that we'll discuss more later).

Rather, narratively-based emotions are an integral & incredibly *useful* (in very practical terms) element of human cognition & decision-making. So, take *that* logic, you *need* emotions—without those little fellas, you ain't nothin'. (You can examine a visual depiction of the above-described cognitive loop by exploring our *Rudimentary Map of Human Consciousness.*)

Mitigating Factors & Complex Emotions

Of course, our endangered-yummy scenario only depicts the most basic of emotions: pain & pleasure. This is mostly because I conveniently kept our scenario free of any real *mitigating factors*. In other words, our scenario involved very simple causal elements (our own accidental carelessness led to a potential loss, quick action resulted in a gain) and highly predictable results (fire will burn me briefly & harmlessly, eating the yummy will give me pleasure).

But life is usually *full* of mitigating factors. *I was going to give half to my starving child. I already lost one hand in a fire. I think I saw the wanderer poop in the campfire earlier.* These mitigating factors can makes us feel *all sorts* of things. (These kinds of context-based emotional mechanics are at the foundation of the currently most-accepted approach to emotions: *appraisal theory*, which provides the basis for emotional models by leading theorists like the late Richard Lazarus & Robert Plutchik. [4, 5])

In essence, each mitigating factor becomes an additional variable in the overall narrative's emotional equation. And these variables—which lead to more complex emotions—are primarily the results of three basic types of narrative judgements: judgements that measure the *validity* (reliability and/or likelihood) of a value loss/gain prediction, *potential* loss/gain, and judgements that measure *other individuals' roles* in a value loss/gain. Which is a mouthful. So before you go back to reread that, let's move quickly to an example...

Since we're going to further torture our poor idiot, let's at least give him a name; we'll call him Rodney (since that's what the R. in R. Salvador stands for, no sense in offending other name-holders). In our new endangered-yummy example, let's say Rodney was joined by the wanderer before dropping his yummy into the fire. In addition (because I can't help myself) *Rodney thinks he saw the wanderer poop in the campfire earlier.* Thus, a mitigating factor has just been added to his "reach into the fire" narrative. In essence, the *validity* (or likelihood) of our story #1's happy ending has been been undercut by the possibility that the yummy has been contaminated by poop. (For the sake of simplicity here, we'll ignore extra narrative branches that might involve Rodney trying to ascertain more clearly whether or not there is actually poop in the fire, and assume he only has his brief distant view of a squatting wanderer as proof. Adding these branches would make the equation more complex, but not illustrate any additional mechanics.)

When compared against the happy-ending narrative, this new poopy-yummy narrative branch seems equally possible. Rodney *wants* the gain of recovering his yummy, but no longer has full confidence in his happy-ending narrative. The result is a different kind of pain-related emotion: *anxiety*. This anxiety is a negative validity judgement. It says *this thing we're about to do or thinking about likely doing because it has a big potential gain, we now doubt to some degree the validity (or likelihood) of that prediction being correct.* And this emotion has a purpose: it wants us to hesitate. It wants to give our brain a few more moments to run new prediction subroutines and determine more possible solutions. It wants a little more time to work its looping thought-iteration magic in hopes of discovering a preferred high-validity happy-ending narrative.

The specific level of anxiety is determined by the phrase we used earlier: *we now doubt to some degree*. The degree of doubt you have is equivalent to the level of anxiety produced—high doubt (low validity/likelihood) means high anxiety (more intense anxiety-related pain). And although anxiety is the product of a more complex judgement, its ultimate result is still to contribute to that core binary emotional response: inhibit or encourage. *Because Rodney was worried that his yummy might be poopy, he felt nervous and hesitated before reaching into the fire.*

The Fear of Losing Yummy

The thing about complex emotions is that they are…*complex*. And in the case of anxiety, it's usually accompanied by another pain-related emotion—one that contributes to the ultimate level of inhibition or hesitation generated when you're worried that your yummy might be poopy. That emotion: *fear*. Although it involves prediction, fear (unlike anxiety) isn't primarily about validity, it's about value—specifically, a value loss. More specifically, it's about a *potential* value loss.

When your brain begins to have anxiety about a desired prediction failing, it's interested in what that failure is going to cost. *If I retrieve a poopy yummy, what's the loss?* (And when your brain is feeling predictive *confidence* about a desired gain, it's interested in how *excited* you should be about that upcoming potential gain.) In Rodney's case, he's calculating a few potential loss scenarios. If he doesn't reach in, he loses the yummy. If he reaches in and the yummy is poopy, he loses the yummy, suffers a small burn, and risks a poop-contaminated hand. If he reaches in and the yummy is okay, he only suffers a small burn. In reality, he only has two choices: reach in or don't. Both predictive narratives produce some fear over potential losses, but because one of the choices (reaching) offers a 50/50 big loss potential and the other (not reaching) a 100% big loss potential, the latter choice produces more fear.

This fear of the 100% loss pushes Rodney toward reaching in (he is *afraid* not to, thus inhibiting any attempt to *resist* reaching). And yet, as he reaches, the anxiety from the uncertainty over the yummy's cleanliness still makes him hesitate momentarily, and possibly experience with it a little more fear over the 50/50 potential loss posed by possible poopiness. These emotions serve the same purpose: to slow Rodney down, just a little, just in case that time can provide him with a unique and preferable solution. But the clock is always ticking. And possibilities like the yummy getting burnt and the wanderer snagging it for himself place a deadline on our calculations. In this case—if he really desires that yummy—even after the fear- & anxiety-produced hesitation, when that deadline comes, Rodney's brain is likely to roll the dice and gamble that it's better to reach than not to reach, poop be damned.

And this is what I meant when I said that emotions are pure logic cut with a dose of gambling. We set up a narrative's emotional equations, add all of the mitigating factors, fill in all of the value-and validity-based variables (determined by previous emotional tags, narrative judgements, and prediction pattern comparison), and then create a final emotional mix intended to guide us toward taking a chance on the choice that seems most likely to achieve the largest gain and/or avoid the largest loss. In Rodney's case, in addition to seeking the largest gain, he's also risking the largest loss: a poopy yummy, plus a little burnt skin & maybe even a poop-contaminated hand. Pure logic with a dose of gambling.

And there's a reason that different emotions are used to measure *potential loss/gain* & *prediction validity* judgements: the combination helps to calibrate our overall inhibit/encourage behavioral response appropriately according to the specific situation. Therefore, if we have high doubt (a low validity judgement) but the potential loss is very small and the action still provides the possibility for a desired gain, the small potential loss lessens the overall anxiety/fear-produced inhibition—making us more likely to take that doubted action.

This is the basic emotional equation that's at work when we do something like spend $2 on a carnival game that we know is rigged for us to lose, yet still *might* win us that cute stuffed bear (and give us that simple feeling of pleasure from having defeated the challenge). When you play the game, you probably feel a little bit of that anxiety over the unlikeliness of succeeding (weakened by the small potential loss) combined with the excitement over the unlikely-but-desired potential gain (an excitement, frankly, that often seems disproportionate to the gain of a mere *stuffed bear*, but we humans are pretty good at *overvaluing* our simple pleasures).

Conversely, even if our doubt isn't very powerful (in cases like a "medium" a validity judgement) but the potential loss is very high, our overall anxiety/fear-produced inhibition is still likely to be fairly significant. In other words, we're *pretty sure* this is going to work out, but the potential loss might be so great that *pretty sure* just isn't good enough. This means we're more likely to hesitate before this action—in the hopes of coming up with something more certain than *pretty sure*. In all of these kinds of situations, our brains are combining the differently-measured emotions of *anxiety/confidence* & *fear/excitement* to properly calibrate our behavioral response using situation-specific calculations that separately account for likelihood & potential loss/gain.

These categories of predictive emotional judgements are central to Kahneman's *Prospect Theory* equations, which show how human brains make these types of decisions by calculating value and probability of predicted results. Kahneman's "value" is our theory's *gains & losses* (measured according to *importance*, *relevance* & *novelty*) and his "probability" is equal to our theory's *validity*, which we actually view as a combination of the *likelihood* of a prediction & *reliability* of prediction data. This *reliability* judgement might be thought of as how much we *trust* the predictive data and/or its source, which can be impacted by factors like sharing a source's beliefs or having a close bond with the source (both discussed later).

But wait...there's more! Complex emotions are not only complex, they're everywhere. And there are still a few emotional complexities to iron out in our *Rodney drops a yummy into a possibly poopy fire* scenario. Earlier, I'd said that there were primarily three basic types of narrative judgements that lead to complex emotions: those that measure prediction validity (anxiety/confidence), those that measure potential loss/gain (fear/excitement) and those that measure *other individuals' roles in a value loss/gain*. My shorthand for these types of individuals: Agents of Value (gain or loss). A teacher, who can potentially confer knowledge value, might be seen as an Agent of Gain. A thief, who can potentially cause you asset damage, might be seen as an Agent of Loss.

We can also make more subtle—and in many ways more unconscious—judgements that lead us to view others as Agents of Gain or Loss: similarities or differences in visual appearance, common or conflicting social/cultural identity, even synchronous or asynchronous physical movements can impact these judgements of other individuals (as demonstrated in research by Northeastern University Professor of Psychology, David DeSteno [6]). These less narrative & more *reflexive* Agent of Gain/Loss judgements are likely tied to mammals' most primitive, least cognitively-based judgements of fellow species-members.

Returning to our contextually/narratively-based emotions: when we perceive someone as a *known* Agent of Gain or Loss (based on a specific experienced or studied act/behavior) or a *potential* Agent of Gain or Loss (based on patterns predicting future acts/behavior) we have different specific feelings toward them. In response to a known Agent of Gain, we feel *gratitude*. *Rodney offered the wanderer half his yummy, and the wanderer felt a good feeling toward Rodney that he could only describe as gratitude.* This pleasure associates that Agent of Gain with memory data that has been tagged as positive.

As we've pointed out, every emotion is a Yin & Yang spectrum. And gratitude's Yang is *anger*—the response to the thief, the known Agent of Loss. *When Rodney retrieved his yummy and saw it was poopy, he felt angry toward the wanderer because he'd cost him the chance to save his yummy*.

The Power of Love & Hate

Gratitude and anger are primarily value propositions. The larger the gain or loss, the greater the gratitude or anger toward the Agent of Value. In more complicated scenarios, level of culpability and/or certainty over culpability can affect the level of emotion generated, but even in these cases, gratitude & anger are still used mainly to reflect value. If the wanderer was starving (increasing the yummy's value) he might've felt more powerful gratitude toward Rodney for sharing. If Rodney was starving, he might've punched the stranger for pooping in his fire.

In contrast to this, *potential* Agents of Gain/Loss are judged using both value *and* validity criteria, because it's about predicting the likelihood that this person will be a *future* Agent of Value. Thus, the emotions produced are slightly different. A potential Agent of Gain triggers *affection*, an emotion so powerful that at its highest level it is basically love. Potential Agents of Loss evoke *animosity*, which can grow into viciously-powerful hate.

One of the things that makes these emotions so powerful is the way they combine a value judgement with a prediction assessment. Consider that the likelihood of potential future losses caused by someone is increased by the number of actual or perceived losses caused by them in the past. So by the time we have predictive confidence in someone's potential to cause future losses, we've possibly already accumulated a good store of strongly imprinted ill-will toward them—which is only increased by the losses we predict they will cause. In this way, it seems that both animosity & affection can grow in a compound fashion.

And yet, because animosity & affection are about *potential* loss/gain, we don't need any *actual* past loss/gain experiences with an individual (or entity) to feel either of these emotions. We just need to believe the individual/entity is capable & likely to cause us future losses or gains. *After your 2-minute conversation with your daughter's arrogant, dumb & clearly-reckless brand-new boyfriend, you despise him. You can feel it in your bones, and you didn't even know he existed 3 minutes ago.* There's still a value judgement here: because the potential harm involves your daughter (very high value) the animosity is more intense than if the guy was just your neighbor's kid's friend. But that value judgement is not based on any previous losses caused by the new boyfriend, demonstrating that these emotions are about *potential* events—and that love & hate can quickly grow from nothing.

The difference between anger/gratitude (*known* Agents) & animosity/affection (*potential* Agents) becomes clearer when we realize than we can both feel gratitude toward someone and still have continued animosity toward them. (Or feel anger, yet continued affection.) Imagine that a homeless person is handed a free meal. They feel momentary gratitude toward the *known* Agent of Gain—this occurs almost no matter who the Agent is (as long as they suspect no malicious intent in the act).

Now imagine that a homeless person has been given a free meal by a well-intended congressman who has led the charge against—and will likely continue to oppose—robust homeless services (and this is understood by the homeless person). The homeless person might still experience some momentary gratitude for the specific act, but they could maintain a general animosity toward the congressman because he is a *potential* Agent of Loss. Conversely, when your spouse crashes the car for the third time and sends your insurance skyrocketing, you may feel some very certain anger toward them in the moment, but nearly simultaneously—or close on that anger's heels—you should (hopefully) be able to look into their eyes and still feel a good measure of affection because of their future potential as a high value Agent of Gain (which is, I know, an *awfully* romantic way to view love).

And to add *even less* romance to the matter of *romance*, I'll share our theory's own special analogy for love's harrowing journey: a two-stage hormone-&-neurotransmitter-driven rocket that sends into orbit a highly-volatile satellite whose speed & trajectory are subject to near-constant (& often *orbit-dooming*) changes. Our *Stage 1 Rocket*—the Saturn V-esque monster that possesses the power to achieve escape-velocity—is that initial rush of attraction, lust (& let's be honest: *obsession*) that accompanies those earliest months.

As the fuel from this beautiful monster wanes & its engines are shed, our more modest (but vital-to-achieving-orbit) *Stage 2 Rocket*—that less-lusty-but-still-intense period of bonding & attachment —takes over propulsion. This is a period that *once upon a time* was meant to result as-soon-as-possible in child-rearing, but these days is just as likely to result in a decision to begin seeking therapy—either jointly or secretly on your own at first. ("Secretly" being an observation that's more *Woody Allen-supported* than *Daniel Kahneman-supported.*)

Once all of *that* fuel supply is spent—then, if final thrusters like *procreation & therapy* have maintained altitude, we at last reach our highly-volatile orbit whose speed & trajectory are subject to near-constant change. And on a week-to-week, month-to-month & year-to-year basis, that orbit is mostly defined by that oh-so-unromantic neural judgement: *whether or not you're able to look into their eyes and still feel a good* *measure of affection because of their future potential as a high value Agent of Gain.* In addition to the effect of ongoing primal breeding cycles—*while the urges last*.

Here again—once our hearts have reached the orbit of affection or animosity—we see specific emotions that are the result of complex judgements, but whose ultimate purpose is to generate that core binary response: encourage or inhibit. Affection draws us to people who can provide us good things in the future (emotional, financial, or parenting support, motivation, knowledge, anything that an individual values) and animosity makes us wary of those who might bring us some sort of harm.

Each emotion reflects our judgement of an Agent of Value and guides our behavior toward them. And every time we gain more value from someone whom we already have great affection for, it reinforces that view of them as a future Agent of Gain, strengthening the affection. This same mechanic is at work with animosity, which is why people often despise an initially disliked President *even more* by the time he's left office. *You thought you hated him when he got elected, but after piling on four additional years of painful, highly-important, highly-relevant, anger-inducing experiences, you can barely stand the guy.*

This *known vs. potential* mechanic also helps explain the roots of the dysfunction that can result in something like an abused spouse continuing to show affection for their abuser. When our brains make predictions about what value we can potentially gain from an individual, many factors are involved. One of the most significant factors is our beliefs—which we'll explore in detail a bit later.

If (through a lifetime of dysfunctionally-arrive-at evidence) I have grown to believe that I am difficult to love, and then (through my limited options) I view this individual as one of my few opportunities to achieve that love, I may be prone to angrily submit to multiple loss-inducing events while still seemingly illogically continuing to exhibit genuine affection toward this individual. This is because I believe they are a novel potential source for something I desperately seek. (I also believe that this kind of prolonged emotional dysfunction eventually "rewires" our emotional responses in a way that we typically perceive as "abnormal" behavior like staying with an abuser.)

One other thing to keep in mind here: under more "normal" circumstances, there are essentially two ways that past experience can help you accumulate enough evidence to result in strong affection or animosity. You can have a high number of small or medium gain experiences that cumulatively provide enough evidence for the brain to judge the individual as a strong potential Agent of Value. Or you can have a smaller number of high gain experiences that provide the necessary evidence. So, even though your neighbor does plenty of nice little things for you month after month, year after year, you still might have less total affection for them than someone whom you only interacted with a few times, but one of those times they saved your life.

The Essence of a Moment

When we mix these judgements gauging matters such as known & potential gain/loss, prediction validity, and known & potential Agents of Gain/Loss, we begin to see the complex chemistry of emotions that define each moment of experience. Consider that all of the scenarios we've dissected thus far are relatively basic narratives. In reality, our constantly-shifting attention, data-rich environment & complicated lives generate a rapid, steady stream of complex interweaving, interchanging narratives. And in any moment we might be surrounded by a diverse collection of individuals about whom we feel a variety of ways. (And, via *empathy*, we might even feel an echo of some other individual's *own* emotions. Empathy also contributes to the emotions evoked by literature & art—the subject of my Story Theory essay.)

Every day is an endless stream of encounters & narratives running the loop through our consciousness, perpetually evoking & generating their own unique emotional results. In addition, the emotional tableau of any moment is likely enhanced by non-narrative emotions that are caused by quick-hit, environmentally-triggered memory pings that evoke associated feelings. *You see a blue uniball pen leaking ink from the cap; it's exactly like the leaking pen your girlfriend handed you after she dumped you.* Here—because the emotions have been encoded into the memory data that has been pinged—the sight of the pen briefly triggers an echo of the pain from that first pen moment.

There are also purely physically-evoked feelings—produced artificially via drugs, or purposefully through injury, activity (like sex & exercise) & urges (like hunger), or mistakenly due to brain or nerve disfunction, etc. Another source of these more reflexive, non-narrative emotions are the primal, pre-programmed genetic responses to specific environmental stimuli: fear caused by the sight of creepy-crawlies, disgust evoked by the taste of rotten food or foul scents, attraction to symmetry in patterns & faces, etc. We also *feel* (although not in an emotional sense) all of those tactile & physical sensations (*smooth or hot*—even sensations like *speed &*

force) which can be perceived specifically or peripherally depending on our attention. Like memory-triggered emotions, these reflexive emotions & physical sensations can all make a similar kind of non-narrative contribution to the *feeling* of a moment. (And although the feelings & thoughts they generate are used in narrative emotional equations—contributing to choices like drug-seeking behavior or Rodney's decision to risk a small burn—they are not essentially a *product* of our consciousness' narrative mechanisms, so we won't discuss them in detail here.)

This wash of widely-varied emotions—each felt in differing intensity, and each derived from different past, present or potential sources—this tableau (combined with those other more reflexive sources) is the essential *feeling* of any given moment of existence. While our consciousness is drawing our attention to data in our environment (& ourselves) and running related internal dialogue narratives, these combined mechanics are also helping to generate the accompanying emotions, feelings & sensations of the moment, which contribute to the overall purpose of our consciousness: to predict results and make decisions, lots of them, every second of every day.

This mix of feelings composing the experience of a moment is roughly equivalent to what philosophers have long referred to as *qualia*—a word that seems to exist only because we had no more precise terminology. But now we have more precise terminology, so let us never speak of that oft-debated, oft-misrepresented term *qualia* ever again!

...Or we'll never speak of it again *after a few more paragraphs*. Before ditching the term entirely, we *should* probably specifically address one very common misperception (or misrepresentation) of "qualia"—one that many over-thinking philosophy-types like to use to prop the door open for the possibility of some *ineffable*, *non-physically-based* quality of mind. This misperception is that there is, for example, some intrinsic & specific qualia-like "sensation" that partially defines (or is the foundation of) our *experience* of something like seeing the color red. This *floaty mind argument* (which is my view of it, not how they describe it) claims that this "sensation of red" is a type of qualia that cannot "merely" be ascribed to the physical processes within our brain —which is, of course, *nonsense*.

The "color red" is specific visual data that we have been taught to linguistically define as the word "red"—a linguistic tag that our culture has kept powerfully consistent for many millennia. For any individual, this word & its associated visual data appear (separately or together) within innumerable personal & emotionally-impactful experiences, and play widely-varying roles in those experiences. In addition, we have been culturally taught to associate that word & visual data with specific ideas & actions (e.g., *red means stop*).

Thus, if you are shown a big red wall & asked how the color makes you *feel*, your response will ultimately describe some emergent combination of the result of all those other (differently-weighted & emotionally-varied) previous associations. There is no innate *sensation* or *feeling* of red that we either all share or that is individually intrinsically & consistently the "sensation of red" to us. (In other words, you likely don't share the exact same *feeling* of red with another version of yourself from a much different period in your life.)

The *sensation* or *feeling* of perceiving or imagining any particular color or object or memory or idea—the feeling of *anything & everything*—is a result of all those types of in-the-moment emotions & memory-associated emotions (& physical sensations) that we're discussing here. Feelings that are (or were) attached to current & previously-stored versions of our sensory or linguistic data via experience.

When we consider the likely complexity of the "emotional fingerprint" created by any moment's mix of varying emotions at varying intensities, we can see why our experiences and memories are capable of evoking such "moment-specific" feelings—which can be both very intense, and in a way indescribable. How *could* we truly describe the mix of feelings that composes a moment? Usually, we pick out the most prominent note among the cacophony of emotions and define the moment that way, reducing it to one of the more basic tags. *I was so…happy. It felt, I don't know, just… depressing. All I can tell you is…I was scared.*

If we were being accurate, we might say something more like: *Well, I was mildly nervous about the upcoming interview, but fairly confident and excited about my date afterward, very annoyed by the gnats in my face, a little scared when I saw that guy who I thought was Joe, and thrilled that the check I was opening was twice what I was expecting!* And keep in mind: that description only included the net emotional results of the different narrative threads mentioned. To arrive at those results, our brain had to provide that other set of sub-calculated emotional values & judgements to be plugged into the main emotional equation (like deriving slight "pattern-pleasure" from the cloud of gnats' visual presentation despite an overall judgement of them as annoying).

In light of all this, it's not hard to believe that the feeling of each moment—its emotional essence—is like those mythical no-two-are-ever-alike snowflakes. It's the most torturous quality of nostalgia: that we seek to recreate the emotional essence of a moment or experience, but in reality, that is nearly impossible.

The Spectrum: Perform or Survive

One of the *coolest* things about the human brain is its capacity to achieve this kind of extraordinary emotional complexity through a system that is, in its own way, extraordinary in its simplicity, its *elegance*. And emotion's ability to create this complexity out of simplicity is akin to the way a wide array of colors can be achieved through different combinations of the 3 primary colors in varying intensity. But instead of having merely 3 colors, Narrative Complexity hypothesizes that our brain's emotional palette has at least 26 "primary colors" at its disposal (13 Yin & Yang pairings)—all of which can be mixed in at anywhere from 1% to 100% intensity.

Now, I know that since I just offered up the number 26, you want to know what they all are—and I promise we'll get to that, but before we do, let's lay out a few more things about our *magical 26*. First, this encourage/inhibit instruction does more than simply tell us to act or not act, it seems to calibrate an entire set of responses—both physical and mental—that better prepare us to confront whatever challenge we face. Before (or as) our brain urges us toward an *action*, it seeks to calibrate our *behavior* prior to that action in a way that gives us the best chance to achieve a desired result.

Therefore, when our brain is flooded with pleasure-based (encouraging) or pain-based (inhibiting) emotions, the emotions are *preparing us to act* in addition to helping us *choose* to act (or not act). Some of these reflexively-triggered "behavioral preparations" or responses specifically differ in response to different emotional combinations (e.g., the reflexive facial expressions & bodily responses that accompany our various emotions). However, according to our theory, there is also a dichotomized set of more neurally-generalized & emotionally-universal brain states that are triggered depending upon which side of the pleasure/pain (positive/negative) spectrum the emotion falls.

In the case of pleasure or encouragement, the positive emotions help to create a "performance mode" in our minds and bodies. This might also be thought of as an "open" state in which we are free to act with more fluidity and greater resource-focus on the task at hand. Basically, the brain is saying *we can be in performance mode here, which requires a devotion of our primary physical & mental resources to this task*. The brain arrives at this decision through emotional equations that determine: 1) this task is worth it, and 2) we can safely devote our resources to this task without exposing ourselves to unnecessary risk by temporarily ignoring other needs (*aka, non-mission-critical neural resource-requests*). We're also prone to devote these resources even if it isn't actually safe, but the action is of such high priority that we're willing to take that risk—which we've probably convinced ourselves is avoidable.

The opposite occurs when our brain is flooded with inhibiting emotions. As opposed to performance mode, our brain and body go into "survival mode." This kind of behavior is reflected in the hesitation caused by fear and anxiety. Instead of creating an "open" (higher performance/ higher risk) state, the negative emotions create a "guarded" state that sacrifices fluidity & goal-focused resource-devotion in favor of caution, protection & more diffuse resource-devotion. Via resource-use that's spread more diffusely to all of our internal & external sensory mechanisms, we are hyper-aware of & ready-to-defend against *any* possibly danger-predicting data in our environment or ourselves *in addition* to focusing some of those resources on the perceived potential loss.

To best understand this dichotomy, it is most useful to examine it at its extremes. Ultimate performance mode is reflected by athletes who are "in the zone" and perform with such fluid physical & decision-making precision that it seems almost inhuman. In this case, all of the positive emotions—pleasure from the accumulating success, growing confidence from their belief in their skills to achieve their goals, pride from the social status gained by their performance —this flood of positivity merges with their actual skill & ability to create a nearly-ideal performance state in which everything else drops away from their consciousness and all resources are freely devoted to their athletic task. They have become the perfect machine for this particular moment.

And when we are in these "hyper-positive" neural states, the way in which these more-focused neural resources are used is likely dependent upon the momentary requirements of the specific task & where we are devoting most of our attention in that moment. Thus, when a musician enters this kind of brain mode while performing, it's likely that their resource-focus will mostly be devoted to their auditory systems —creating a heightened, more vivid & detailed auditory experience, which aids in their musical performance (and depending on the instrument, there might also be heightened tactile or physical responses).

Later, such a musician might be able to describe the performance in extraordinary detail—while having little memory of specific visual data, like the actions of the crowd. Except, for example, in those moments *after* they complete a song or performance and their brain (still in its hyper-positive state) turns its attention toward the cheering crowd. Now that those extra-focused resources are no longer needed by the auditory & physical systems, they can be used for the primary task in this moment: looking over the crowd. This helps to create a momentarily extra-vivid & detailed visual experience as they take in the full sweep of their adoring throngs.

This kind of shift-in-focus/shift-in-resource-devotion is also reflected in the way many athletes describe those in-the-zone experiences. For example, when standing at the plate before a big moment, baseball players often describe the vividness of the crowd & the sea of flashing bulbs. But once the pitcher winds up, that same player often describes losing all sense of the crowd. With their extra-focused resources now devoted to hitting, the sight of the rapidly-oncoming ball fills their visual field with extraordinary detail—as is frequently stated: *they can see the seams on the baseball.*

At the other end of this spectrum is paralyzing fear—those moments in which all choices seemingly lead to great loss or harm, making you so afraid that you are literally frozen, unable to act at all. And in your frozen state you feel almost *animalistic*: nearly wordless, cowering, trembling, eyes darting frantically between each rustle of sight & sound, ready to protect ourselves, to lash out violently if provoked. In these cases, your brain isn't interested in what your consciousness might want to focus its resources on—you cannot afford to leave any aspect of this moment fully-unattended. And your brain doesn't *want* you to fully-focus on any specific task right now—it's trying to *inhibit* your actions until it knows it's safe to "un-guard" itself. This is an extreme response to the same impulse that made Rodney hesitate before reaching into the fire for his possibly-poopy yummy.

In these fearful or guarded neural states, we naturally still retain some primary focus on the identified threat or loss but—because the diffuse resource-distribution limits resource-use by any specific system—that focus (e.g., visual resources devoted to a threat) is likely much more narrow than the rich, broad focus experienced in positive neural states. Thus (returning again to those oncoming baseballs) when a timid Little Leaguer returns to the plate—after being hit by a pitch his first time at bat—and fearfully stares down yet *another* baseball speeding toward his helmet, he will very likely have a strong-but-narrow visual focus on the incoming projectile. Nonetheless, in its resource-deprived state this visual focus *does not* result in any capacity to *see the seams on the baseball* before it nearly beans him (we'll be kind & assume he learned from his first experience & ducked out of the way this time).

Ultimately, when we experience something like paralyzing fear or anxiety, your brain is begging you to wait until you can find some solution that doesn't involve a major loss. *Don't move. At all. And keep an eye on that, but stay alert! If you notice* anything—*protect yourself! Let's see if we can figure something out before you do this thing that is very likely to end very badly.*

Performance & Survival. Open & Guarded. Encourage & Inhibit. Pleasure & Pain. Yin & Yang. This is the spectrum upon which all emotions are measured & expressed. In the end, we're simple creatures—it just takes a whole lot of calculating to get there.

The Purposes: Imprint & Signal

Yes, I know, *what about the magical 26?* Getting closer…*promise*. But there's a distinction within our emotions—one we've already acknowledged—that I want to bring to the forefront before revealing the 26. It's the distinction between our emotions' two basic purposes: *imprinting* & *signaling*.

"Imprinting" is the encoding of data with a particular positive or negative value at the time of incident ("*Ouch! That red glowing stuff is hot*" or "*Mmm! This stuff I'm eating is yummy*"). As we'll discuss in Essay #4, this emotional imprinting also plays a key role in how weakly or powerfully an experience is *remembered*. The greater the *intensity* of the pleasure- or pain-based emotion (likely determined primarily by the overall gain/loss value of the event) the more strongly the event is imprinted into your memory.

Our theory also hypothesizes that our imprinting or "tagging" process works slightly differently when we make judgements about other entities (individuals, groups). In this process, emotions can both help encode the entity itself with a value, and help create/strengthen a connection between the entity and other data that has been encoded with a value (i.e., a gain/loss event). This is the mechanic that allows us to associate anger-generated, negative-value data with someone whom we actually have affection for—*without* changing our overall perception of them as an Agent of Gain. (In Essay #4, we'll discuss more deeply how this process is managed.)

Our emotions' other purpose is "signaling" or prompting, which is the primary emotional mechanic we have been discussing thus far—guiding our actions & behavior toward a desirable result. Although most of the signaling examples I've provided have been fairly straight-forward (e.g., *fear* signals behavior that helps mitigate a potential loss), our full matrix of emotions will also detail some of the more complex behaviors that our emotions can signal. These are the most sophisticated of our primary emotional pairs (and might be the most recent to evolve, which we'll discuss more later).

Some emotions are likely more heavily-weighted toward either imprinting or signaling, depending upon the kind of judgment they are designed to make. For example, pain—which is the result of an *actual* loss, and therefore a reliable indicator that this action will also be harmful in the future—is likely a stronger *imprinter* than fear. This is because fear is triggered by a *potential* loss, and is thus more-likely geared toward *signaling* (prompting) behavior & actions that help us to avoid or mitigate the *not-yet-happened* loss. In fact, if your fear is effective-enough in helping you to actually find a way to avoid that loss, then your brain would probably find it more beneficial to imprint the experience more *positively* than negatively.

Thus, it would make sense that fear's imprinting power be weak enough to be *out-imprinted* by emotions that actually judge whether the experience was ultimately positive or negative. Similarly, it's likely that emotions reflecting *actual* prediction success or failure—affirmation & surprise—are stronger imprinters than primarily *signaling* emotions that reflect *potential* prediction success or failure—confidence & anxiety.

There is also another kind of "signaling" purpose that our emotions serve—a kind of signaling that we noted when describing those more specific reflexive physical responses generated by different emotions: facial expressions. The widely-varied facial expressions (& accompanying "body language") generated by different emotions play a key role in *expressing* or *communicating* how we feel—both to other people and to *ourselves*. (And although the basic templates for pain-based & pleasure-based facial expressions are likely inborn—aiding infants & toddlers in their early attempts to express & to comprehend expression—recent research has shown that our emotion-based facial responses are also deeply influenced by learned cultural cues.)

In the case of *ourselves*, there is a kind of internal "feedback loop" that can result from reflexive physical emotional responses like smiling when experiencing some gain or positive result: the physical act of smiling seems to enhance (or help to perpetuate) those positive feelings that triggered the smile. This kind of feedback loop likely helps us to sustain those "preparatory" emotional states (and thus sustain the situationally-advantageous neural-state) that precede actual decisions & actions *without* having to continually cognitively *re-assess* the situation in order to help continually "re-trigger" that situationally-advantageous neural-state.

In the case of expressing these emotions to other people, there are obviously myriad powerful communicative & social advantages provided by the capacity to visually demonstrate & identify various emotional states. From a cooperative & knowledge-sharing perspective, instantaneously perceiving a companion's expressed emotional response to stimuli that is novel to you—but familiar to them—is an almost-magical & wordless way by which that companion can communicate (& allow you to make personal use of) data derived from their own experience. This ability also allows you to wordlessly (& sometimes distantly) detect things like whether or not that companion is expressing their dire need of your help.

From an adversarial social perspective, instantaneously perceiving, for example, a possible enemy's expressed emotional response to *you* can obviously be extremely useful in helping you to quickly take any survival-aiding actions before that survival is actually in jeopardy. Indeed, visually expressing or assessing everything from fear to confidence to guilt can aid in effectively choosing how to manipulate or respond to social conflicts.

Despite their varied purposes & applications, all of these imprinting & signaling mechanisms play a vital role in calculating & enacting the results of our brain's emotional equations. Imprinting allows memory-based data to have actual values when plugged into those equations, and signaling ensures that the results of the equations guide our behavior, actions & neural/physical states in useful or advantageous ways based on the known data. (Narrative Complexity's layered, multifaceted view of our emotions' myriad & interconnected functions reflects the kind of non-exclusive & integrated approach to emotional function suggested in the 2013 paper by Farb, Chapman & Anderson, *Emotion: Form Follows Function.*[7])

Part of what makes this system plausible is the fact that all decisions & emotions are data-based. Not only data-based, but data-based in a way that is ultimately binary, which is the way our brain primarily functions. In the end, everything in the brain essentially comes down to an unimaginably vast array of on/off switches. Emotions make maximum use of those switches. *Complexity from elegance.* If you could use only three words to describe how the human brain functions, those would be the three words.

The Secrets of Beliefs

So, yes, we've *almost* arrived at that part. The part where we reveal *The Mothership of Emotions*. But there's just *one little concept* that I need to slip into your brain before we visit *The Mothership*. Actually, it's a pretty big concept, one that might be the most powerful force in shaping our most important decisions: *beliefs*.

According to our theory, there are special emotional pairs that are specifically designed to use our beliefs to generate feelings. And these beliefs provide the foundation for a vast number of the decisions we make. *You believe in God. You believe in the principles of conservatism or* *liberalism. You believe that love is always good and violence is always bad. You believe violence is a necessary evil.* If you were to catalog them, your list of personal beliefs might seem nearly endless. Yet, the list would still have an hierarchy. And if a decision pits two opposing beliefs against each other, the stronger belief is very likely to win out. So what does that mean, for a belief to be *stronger* than another? To answer that question, we first need to answer a more fundamental one: *what is a belief?*

In the view of Narrative Complexity, a belief is, in essence, a high-value, high-validity prediction trope. It expresses a basic (although often complexly arrived at or applied), important, broadly-applicable and over-arching prediction that has achieved very high validity through the accumulated experience or study of actual or perceived-to-be-true events. *I believe forgiveness is always better than revenge.* Or more purely: *I believe in forgiveness.* Translated: in any choice that can be reduced to an act of forgiveness or revenge, choosing forgiveness is highly-likely to achieve a more desirable ultimate result.

The higher a belief's related value (e.g., *your soul's eternal survival = extremely high value*) and the higher its validity (*being taught something from the moment your memory began, by people you implicitly trust = very high validity*) the higher a belief rises in the hierarchy (*Above all else, I believe in God*). These top-level tropes are decision-making gladiators—taking on all contradictory ideas or choices and slaying them with the power of their "truth." Who are these gladiators really? Purely-reduced & powerful prediction models that represent something we assess to be both a highly-valid prediction in almost all circumstances & settings, and a prediction that relates to many high-value goals.

Cheating is bad. All success requires hard work. These are superseding predictors, the express lane of decision-making, because if can we find a way to apply this predictive pattern—even without examining related data in detail—we think there is a strong likelihood of goal-success. Which does not make a belief true, it just means you "successfully" applied it or "know" it has been successfully applied enough from your perceived personal experience or your study of "reliable" sources to make it rise to the level of a belief.

And this mechanic reveals the source of many seemingly illogical behaviors and beliefs, which are actually based on very logical choices by our brain—unfortunately, in these cases, our brain has arrived at this logic through bad data or data that has been misinterpreted (often through the application of other powerful, but false beliefs).

For example, long-ago seafarers behaved in all kinds of illogical ways because they believed sailing too far would send them off the edge of the world. This belief was founded on the superseding belief that the world was flat. *This* belief was arrived at through a lifetime of misinterpreted evidence (it *looks* flat, *all the time*) and bad data sources (*everybody* says it's flat). It was almost impossible for those sailors to imagine that the sea wasn't a purely flat and likely finite entity, because they had no "valid" pattern evidence to build a different belief on.

Thus, we have *confirmation bias*—because when we judge contradictory data for validity we often can't even imagine it as true, which makes us more likely to seek out & choose to trust data that reinforces what we already believe.

[Dude from the future speaking to the long-ago seafarer.]

Dude: *Look, trust me, the world is round. That's why you can't see forever along its surface, because the surface is curved!*

Seafarer: *Right. I can't see forever because it's too far away. And on the other side of this "round" world, I suppose they're upside down and still sticking to the ground? Don't think so.*

Dude: *Gravity man. Heavier objects attract smaller, and the earth is huge!*

You can see this conversation isn't going anywhere. To the long-ago seafarer's brain, what the Dude is saying is inherently *not true* and thus, nearly impossible to tag as valid. This also makes it nearly impossible for the Dude's true, but unconvincing evidence to change the ancient mariner's belief. One way to avoid this trap is to make *"Doubting your instinct to believe in something"* one of your highest level beliefs, which is a way to "short circuit" confirmation bias. This belief does that by making doubt supersede certainty, which provides your brain with a logical, high-validity reason to give contradictory data a second look. And this allows your brain to accept this data as valid *despite* the fact that it contradicts what you "know" to be true.

It's an awfully tricky trick—which is why most of us are total suckers for confirmation bias. But the use of this trick is why the scientific method, over time, has been able to initiate major changes in human beliefs: because it is built on *skepticism*—that belief that doubt supersedes certainty. This has helped science-based endeavors to accumulate enough valid evidence and repeatedly produce enough confirming data to slowly change many of our beliefs.

Despite all this, to our brain, confirmation bias is not a flaw. Most humans do not have the luxury of being able to treat all evidence as possibly equal without further, detailed examination. It's much more efficient to build beliefs on accumulated past evidence and trust those assessments, otherwise we might be frozen by the possibilities of what might be the *real* best decision. In fact, using the evidence that we've already gathered is essentially the *only* way we can create our beliefs. Our whole system of consciousness is founded on trusting our original value tags & validity judgements and building upon those. Yes, this means that humanity can get mired in ultimately false beliefs for a long time, but in a way many of these beliefs are *functionally* true. This means that the application of these beliefs still works within the framework of what is *actually* true well-enough to aid in our survival.

In other words, yes, there were great benefits to be had by understanding that the world is, indeed, round. But the belief that it was flat still embodied enough *actual* truths about the world to make it functional. *If we move consistently in one direction, we will arrive at a different place. When we encounter a valley or mountain, it will not continue in perpetual incline or decline, but be surmountable at some point, etc.* These might seem to be uselessly obvious premises to us, but to ancient man these truths were *functionally* more important than the belief that the earth is round, and therefore highly-useful despite contributing to a false belief.

And this appearance of functional truths within an ultimately false belief is not an accident. This occurs exactly *because* our brain is using that time-tested experiential-data-based method to build the belief. *Some* of that belief-building data has been interpreted in valid ways, and is therefore specifically useful even though we've gotten the big picture wrong (which leads to other problems, but nobody's perfect). Thus, confirmation bias has survived, because even though it can divert us to the wrong track, that track can still get us to where we need to go at that moment.

Which is all good & well, but what exactly do these beliefs have to do with emotion? Suffice to say: our brains *do not* like it when we let the lure of big pleasure or big gains usurp the supremacy of our beloved beliefs in the decision-making process. *Sure, this seems awesome right now, but think BIG PICTURE. All the good you can get from this ain't gonna make up for all the bad that's likely right on its heels.* Remember: every time your brain is making you feel terrible, it's just looking out for you. Your brain really *is* in your corner, even when it feels like it isn't.

The Mothership of Emotions

Okay, no more stalling. Following is our Emotion Matrix containing the magical 26 —the 13 base pairs of Yins & Yangs. You are now invited to board: *The Mothership…*

The Mothership of Emotions [Matrix of Primary/Complex Human Emotions]

The Spectrum —> Narrative Triggers	***Performance (Open)*** Emotional Pairs	***Survival (Guarded)***	Primary Purposes
Known Value Gain/Loss	Pleasure	Pain	• Encode data as helpful *or* harmful • Signal behavior that perpetuates gain *or* stops loss
Potential Value Gain/Loss	Excitement	Fear	• Encode data as helpful *or* harmful • Signal behavior that helps ensure gain *or* mitigate loss
Global Value (Known & Potential) Gains/Losses	Happiness	Sadness	• Signal behavior that prepares us to: - expend/risk resources in times of perceived abundance, *or* - conserve/protect resources in times of perceived scarcity
Known Prediction Success/Failure	Affirmation	Surprise	• Encode prediction data as reliable *or* unreliable • Signal behavior continuance *or* cessation
Potential Prediction Success/Failure	Confidence	Anxiety	• Signal behavior that helps ensure prediction success *or* mitigate prediction failure
Known Agent of Gain/Loss	Gratitude	Anger	• Signal behavior toward entity that either: - reflects openness and strengthens bond, *or* - protects against and seeks "restitution" for loss • Associate entity with gain *or* loss data
Potential Agent of Gain/Loss	Affection	Animosity	• Signal behavior toward entity that either: - reflects openness and strengthens bond, *or* - protects against and seeks "restitution" for any previous outstanding losses • Encode entity as helpful *or* harmful

The Mothership of Emotions [Matrix of Primary/Complex Human Emotions]

The Spectrum —> **Narrative Triggers**	***Performance (Open)*** **Emotional Pairs**	***Survival (Guarded)***	**Primary Purposes**
Known Need of Agent of Gain/Loss	Generousness	Selfishness	• Encourage specific act of aiding/sharing with Agent of Gain *or* inhibit specific act of aiding/ sharing with Agent of Loss
Potential Need of Agent of Gain/Loss	Magnanimity	Greed	• Encourage behavior that prepares us to aid/share with Agent of Gain *or* to protect resources from Agent of Loss
Known* or *Potential Social Status Gain/Loss	Pride (In Self)	Embarrassment	• Encode data as "socially" helpful *or* harmful (in terms of prestige in specific community) • Signal behavior that perpetuates gain *or* stops loss
Known Belief Compliance/ Violation (by Other Entity)	Pride (In Other) *[Root of Covetousness]*	Disgust *[Root of Jealousy]*	• Associate entity with "model" *or* "avoid" behavioral data • Signal supportive *or* antagonistic behavior toward entity
Potential Belief Compliance/Violation (by Other Entity)	Admiration *[Root of Envy]*	Disdain *[Root of Resentment]*	• Encode entity as "model" *or* "avoid" • Signal supportive *or* antagonistic behavior toward entity
Known* or *Potential Belief Compliance/Violation (by Self)	Satisfaction	Guilt	• Encourage belief-compliant behavior *or* inhibit belief-violating behavior

[I chose not to include *Engagement/Boredom* because they seem to be a general mental response to the *presence* (engagement) or *absence* (boredom) of useful or novel data in our environment or within whatever we are specifically evaluating. Instead of producing actual pain or pleasure on their own, these "mental states" seem to reflect whether or not there is any possible emotion-producing data present. Thus, *engagement* opens the door to all emotions (which are actually what produce the pain & pleasure, and *keep us* engaged) and *boredom* leads to almost no emotion, a state which makes us want to move on and find something to feel.]

The Mothership's Alien Language

I know, *I know*—you have questions. And complaints. Before you toured *The Mothership*, you were thrilled it had finally arrived (anticipating that value gain). But now that you're aboard, you might be perturbed. *Where is my favorite emotion?! How can you claim this is complete? Magnanimity!? Affirmation?? What the hell!?*

I understand. And don't worry, your favorite emotions haven't gone anywhere. Think of it this way: you're looking at red, blue & yellow, and begging to know why *fuchsia* isn't there. *It's in there.* But we need to work a little alchemy in order to show it to you. And there's something else: *what exactly does fuchsia mean to you?* Sure, we can all eventually agree on what's generally red, blue & yellow—even green, purple & orange. But when we start to get into those *subtle shades* of color & emotion, we also get into that *malleable words* area. Here we begin to see some of the drawbacks of a language that allows for imprecision—a system in which certain words represent less frequently encountered ideas, and are therefore more reliant on specific personal experience for description, as opposed to more cumulatively developed & more culturally reinforced fundamental ideas (fuchsia vs. red).

Nonetheless, before there's a mutiny, let's work a little alchemy and try to make some fuchsia. *Disappointment*. Here we have a combination of surprise (*we thought we were going to ace that test*) and the simple pain of loss (*our failure cost us an "A" in the course*). Conversely, the surprise of an unexpected "A" (prediction failure + value gain) instead creates a feeling we might describe as *delight* (which helps give a positive tag to a gain event that otherwise might've been seen merely as a prediction failure).

But let's return to the disappointed student (because they're more fun to mess with). The student's disappointment might be augmented by other factors. *I should've studied harder* produces *guilt* (they violated their belief: *Success requires hard work*). And when they imagine telling their parents, they begin to experience the inevitable *embarrassment* from public failure (loss of social status). *And* because of their strong affection for their parents (which makes them want, among other things, to be admired by those parents) this failure registers as an even higher value loss, amping up the pain of the embarrassment & guilt to the level of *shame*.

Now imagine that in the back of that student's mind, they suddenly realize that this failure might have the eventual bonus of lowering their parents' expectations, allowing them to imagine future gains in affection achieved at a lower cost (less studying & other success-related effort). Here their brain pumps out a bit of *excitement* over these potential future gains.

In reality, the shame of the moment is probably powerful enough to quell any real feeling of excitement, but its small pleasure still registers—most likely in a way that they perceive as "momentary relief."

When the student saw the unexpected "F" on their test, and realized they'd just lost their "A" in the course, and thought about telling their parents, they were filled with disappointment & shame. Then, for a moment, they imagined a new future in which their parents stopped expecting so much, and felt a small respite from the pain.

Of course, that still might not be *exactly* your description of fuchsia, but we can probably at least agree on which paint matches the curtains now. Keep in mind: it's not so much about the words as it is the judgements they represent, and then tying those judgements to specific pain or pleasure behavioral responses—some of which are more universally recognizable than others.

The less recognizable primary emotions & their sources are, in a way, "camouflaged" because they are rarely felt in total, focused isolation. Consequently, we aren't as compelled or likely to determine their specific narrative triggers (unless, of course, you spend a lot of time in therapy). This means there are some basic emotions that we never really think to distinguish on their own. For example, let's examine that simple (& almost *overly-familiar*) feeling of *affirmation* that you get from positive feedback when playing out a successful predictive pattern. At first glimpse this seems like a pretty flimsy emotion, especially compared to its pair: *surprise*, which is easily (& often powerfully) quantifiable to all of us. But the emotional juice from affirmation is what, for example, video game designers and mystery writers are doling out along the way to get you to the ends of their creations. Every hint revealed along the story's path (confirming the narrative that we are predicting) and every glowing, animated star that pops up en route to the end of a game level (confirming your ongoing success in solving the puzzle)—all of this pleasure says to your brain: *yes, keep going, keep thinking this way.*

And if we look more closely at "unbalanced" pairs like surprise & affirmation—where one half feels more powerful & identifiable —we can see where evolution is likely at work. Surprise *needs* to be more powerful. It's often trying to stop you cold: *woah, that's not what we expected, hold up!* But its pair, affirmation, would probably prefer we stay in the flow of whatever we are (successfully) doing. It just wants to make sure we're positively noting our success along the way. (Here again, our "Guarded vs. Open" mechanic is at work.) Thus, we can see how, over time, these differently weighted usages resulted in differently evolved characteristics within some emotional pairs. (Much research actually suggests that our brain weights almost *all* pain-based emotions more heavily than pleasure-based—something reflected in Kahneman's *Prospect Theory*, which shows that potential losses tend to carry more predictive weight than potential gains in our brain's decision-making calculations.)

We can also see this kind of evolution in *guilt* & *satisfaction* (belief violation & belief compliance). Consider that beliefs are, by definition, already associated with high value & high validity. This makes us generally more likely to comply than not to comply. Thus, *satisfaction* doesn't need to work very hard to reinforce our belief-compliant behavior—our behavior is naturally belief compliant. Satisfaction, like affirmation, is just produced to help keep us going: *excellent*, *you're doing the right thing*, *keep it up*.

This lack of emotional juice when we act belief-compliant is likely one of the reasons why we usually want to tell other people about events such as a our own acts of kindness. Even though we feel some genuine *self-satisfaction* from, say, saving a dog who was hit by a car *(I believe in aiding all creatures in need)*, our satisfaction still might not be as strong as our desire to tell other people—which provides that juicier, more powerful social status reward of pride (something that requires an audience).

Contrary to satisfaction, *guilt* is triggered when a belief's innate power is *not* doing its job—when a belief is being undercut by something like the potential for strong pleasure or big gains (or the desire to avoid a big loss). Thus, guilt has to have some serious juice—because in many cases, it's our last line of defense against a very bad decision. This kind of role likely led our brain to accede to guilt-heavy mutations over the course of evolution.

In this way, we can see how the evolution of emotional pairs is similar to the evolution of more concrete features, like our limbs. Once upon a time, the fins & paws that became limbs were fairly balanced in composition & effect, but as the needs of each end of the mammal grew more specific, the limbs adapted differently (while still remaining fundamentally similar & clearly part of the same original mechanism).

Now, we could continue to scour the emotional spectrum in hopes of eventually hitting everyone's favorite & thus-far-unnamed emotional combo color—but, y'know, *that'd be nuts*. There are way too many hues hidden in the rainbow. However, the colors are all there to mix for yourself. And to show you just how easy (and fun) mixing can be, we'll do *one more* combo color—my own favorite emotion, *melancholy* (the *bittersweet* kind, as opposed to a pure shade of sadness). What I believe most people are describing in these cases of melancholy: the simultaneous experience of pleasure or happiness in response to a *current* moment of value gain combined with the pain or sadness of predicting the *future* loss of the source of your current happiness. In other words, the joy of watching your toddlers play—a current value gain—can be tinged with melancholy if you start to perceive the fact that someday they will no longer be toddlers—a predicted future value loss.

While you're trying to locate your own favorite emotions, keep in mind that some of them are essentially a word that describes a primary emotion in differing intensity: powerful guilt (strong associated loss or violation of a strong belief) is often deemed *remorse*, whereas less powerful guilt might be expressed as simple *regret.* Similarly, *annoyance* is basically a description of very minor *pain* (those bugs in your face cause tiny, but frustrating losses in resources like mental focus). But we're starting to scour the rainbow again, so —*scouring officially ceased*. (You can do a little more scouring at the end of this essay, which lists & describes 14 of the more common "combo-color" emotions.)

Deep Inside The Mothership

Instead of exploring more emotional blends & hues, let's look more closely at a couple of the primary emotional pairs—the ones that seem to need the greatest clarification: *generousness/selfishness* & *magnanimity/greed*. The former pair is easy enough to conceive, but the latter seems almost unnecessary in light of the first. Here again, language complicates matters. In practical terms, humans haven't had much reason to distinguish something like "selfishness" from "greed"—basically, we consider those words synonyms. In both cases the result is the same: *we're keeping it! (or taking it!)*

But our evaluation of another entity's need as *known* (current) or *potential* (future) is necessary to affect the proper kind of behavioral response in each case. If the *yammering homeless guy* on the corner wants money as you walk by him, your momentary *selfishness* might keep you from handing him a buck. But what if you're worried that the government is going to come around *next year* asking for a big income tax hike to help feed *those worthless indigents?* In that case (because you're a greedy jackass who sees the government & homeless people as Agents of Loss) you might actually hide your money in some offshore bank accounts—so when the Feds come asking, it's protected. *Greed rears its ugly head.*

Conversely, if your kid wants money for pizza tonight, you have to be able to distinguish that need from their need to pay for college someday, which requires an entirely different set of actions, behaviors & long-term evaluations—as opposed to one simple act of fulfillment that is primarily dependent on your current resource status.

And these emotions work much like their cousins anger/gratitude & animosity/affection: you can be motivated to feel situation-specific generousness toward an entity that you otherwise generally behave greedily toward & vice versa (e.g., you donate specific disaster relief to a nation that you otherwise support a general embargo against, or you selfishly refuse to anté up for your kid's pizza tonight because you want the cash for beer, but still magnanimously sock away money for their tuition someday). That's because, as similar as these feelings are, they are still the results of slightly different narrative judgements. And this distinction allows your decisions to take into account relevant current & predicted resource status when deciding how to most efficiently & beneficially share or protect your resources when necessary.

Which just leaves us with one more sub-system to examine aboard our Mothership: *covetousness/jealousy* & *envy/resentment*—whose roots are, respectively, *pride(in other)/disgust* & *admiration/disdain*. Since we consider their root pairs to be Complex Emotions, we might think of these other branches as *Very Complex* Emotions. Actually, *all* of our "fuchsias" (like disappointment/delight) are Very Complex Emotions. Which is to say, at first glance they appear to be complex, but primary emotions—until you look a little closer, and realize that all of their component narrative judgements and desired behavioral results can be arrived at through some combination & application of our *magical 26*.

I've specifically noted jealousy, et al, on our emotion matrix (even though they aren't a primary pair) because these are actually among the most powerfully identifiable emotions, and their pairings so mimic the other complex Yins & Yangs that they truly *look* like primary pairs. But jealousy/covetousness & resentment/envy are *very complex* because they involve: judging another entity's belief compliance (pride/disgust), *and* judging a value gain by that entity—a gain that you view somehow as a personal loss, which triggers a combo of pain, generousness/selfishness, and possibly anger or disappointment. (Keep in mind, this "personal loss" doesn't require that you ever really had a chance of having it—to our brains, it's enough to simply want it for yourself & not get it.)

My lazy co-worker (Belief alert! "Success requires hard work") just got the promotion I wanted. I'm pissed. And, frankly, I'm jealous.

Well, Anne got the promotion I wanted. But the truth is she works so hard around here, she deserves it. Still, I'm disappointed. And I really covet her new office—which is terrible, isn't it? I should be happy for her.

It's difficult to be happy for other people when their gain looks like our loss—but when their gain actually reinforces our beliefs, our brain still wants to make sure we find a way to tag the experience positively (thus, covetousness). This is because those actions & behaviors have value to us as an *effective model* of how our beliefs can help us to achieve what we want.

Conversely, when someone else's delicious gain is achieved through behavior that violates of our beliefs, our brain wants to make sure that we still tag this behavior as negative, *despite* the fact that it provides a model for achieving something we might want. So even though you also want that big sailboat your neighbor owns, you don't want to be tempted to set up a Ponzi scheme like he did in order to buy the boat. (Assuming your beliefs predict that the temporary gains from such behavior will likely be followed by dire results.) Thus, jealously gives us the permission to feel negatively about his gain in order to help reinforce future belief compliance (particularly in the face of desired gains like a big sailboat).

Culturally, we tend to view jealousy and covetousness in the same negative light, but this is one of those illogical behaviors based on a learned false belief (one that had logical origins). The roots of the word "to covet" were related to inappropriate sexual desires (this is buried in the word's etymology). But long ago we discovered that the *emotion* of coveting applies to our desire for anything of value that's possessed by someone we respect—even symbolic items, like a job title—which led us to appropriately expand the word's usage. Nonetheless, its original negative association remained, creating the foundation for a false belief: *Coveting is bad.*

The "taboo" of covetousness (taught in ancient religious texts) was originally created by our culture for a good reason. It helped us to avoid a powerful, primal & non-narrative urge: *your neighbor's wife* (whom you might succumb to coveting hands-on, even if you really respect your neighbor, *and* your own wife). But, as we observed, the idea of coveting has long been applied to that whole non-sexual universe of value gains deservedly-achieved by others—gains that we are (usually) much better at controlling our desires for (or at least we're more likely to be deterred by the penalties in place, which your neighbor's spouse is also good at overriding). And these non-sexual gains are the ones that our brain *wants* us to covet—because it knows it can use this data to help us to achieve our own future gains via belief compliance.

When we're jealous, the "ickiness" of the feeling toward the other person comes from our disgust over the belief violation that is at the heart of their value gain. Conversely, your desire to fight your own loss pain in order to "be happy" for the coveted gains of someone you respect—that positive impulse is rooted in your pride in their belief-compliant behavior. So go ahead—*covet* all you like. It's good for you. Just keep your envious eyes (and your hungry hands) off your neighbor's spouse.

A Final Filmstrip: Emotion's Evolution
Alas, the time has come to disembark *The Mothership*, and leave behind all its high-tech, evolutionarily-fancified brain mechanisms. Your own brain, I'm sure, would be happy to take a respite from all those wacky, mind-bending emotional equations. So we will. *Consider the chalkboard cleared.*

But before you go, let me pull the screen down over the board, switch off the lights, and roll out one of those old filmstrip projectors (kids, imagine an ancient PowerPoint presentation with *way better* analog-ish ambience). And don't put your head down on your desk—*you're gonna wanna see this.*

Because our speculation about emotional equations has been based on very familiar experiences & a mathematic *Prospect Theory-supported* approach, its conclusions are in many ways quantifiable. The speculation in our filmstrip, however, is more... *speculative.* Which is, frankly, what one would expect from a story about the evolution of emotion. Nonetheless, the tale is a compelling one. And at the very least, we know that modern human emotions had to come from *somewhere*. And that somewhere is *exactly* where our filmstrip begins...

It is 700 million years before humans ambled onto the evolutionary stage. A little roundworm with an unfortunate first name —Caenorhabditis elegans—is squiggling along in the muck. And little C. elegans has something in common with us: he likes to eat. Not only does he like to eat, he *expresses* this desire using clever devices that we also make heavy use of in the expression of our desires: *neurotransmitters*.

In particular, C. elegans is using serotonin and dopamine, which play significant roles in our own brain mechanics (they are key players in producing & manipulating our pleasure/pain responses). When Mr. Roundworm encounters positive stimuli, like food or a mate, serotonin is released—helping to enact motor scripts like bacteria ingestion. In addition, when his worminess rubs up against that yummy bacteria, dopamine is released. The dopamine helps to inhibit the creature's locomotion motor scripts—slowing him down & allowing him to spend more time in the presence of the food. And if he's really hungry, more serotonin is released—this dose helping to inhibit his locomotion even further, ensuring he eats every last bacterial bite. [8]

I know what you're thinking: *this C. elegans guy sounds like an uncle of mine.* And, yes, from a broad universal perspective, we're not all that different from our wormy planet-mate. But 700 million years is a long time. And our use of these neurotransmitters is so much more diverse & complex than C. elegans' that it's like comparing an abacus with an iPad. Sure, they both calculate stuff with similarly clever efficiency, but an iPad can calculate *a whole lot more stuff*. And not to make C. elegans feel worse about itself, but plenty of tinier & earlier creatures were using neurotransmitters to affect behavior (even lowly paramecium use serotonin when swimming).

Humans didn't evolve from roundworms, but our earliest chordate ancestors (who appeared about 500 million years ago) and roundworms emerged from related evolutionary branches. In fact, scientists have found in C. elegans some of the specific kinds of serotonin receptors that humans use today. [9] And in its simple existence we can see ancient sparks of those relationships between resources (*food*), "feeling" (*neurotransmitters*) & behavior (*stay here*) that are at the root of our complex emotions.

As we said, 700 million years is a long time. And although roundworms hit an evolutionary dead-end, early chordates' simple neurotransmitter- fueled commands *"stay here & eat"* and *"stay here & reproduce"* eventually evolved into early vertebrates' more complexly regulated (but still basic) resources, feeling & behavior relationships. (Thanks to more robust & diverse neural structures & neurotransmitter mechanisms.)

The result was likely a system of primitive *proto-emotional* pairs that helped those early vertebrates to manage: *hunger(thirst)/satiation*, *lust/repulsion* & *strength/fatigue.* Those would cover all of an early creature's basic needs (and later probably composed an average hominin evening in the cave: eat, drink, screw, sleep). *C. elegans politely raises its tail: "Hey, I basically do all of that stuff too!"* Which is true, but more complex creatures began to require resource-acquisition strategies more complicated than *squirm toward that chemical marker & hope I squiggle over something to eat.* Thus, the neurotransmitter-fueled behavior signaled by proto-emotions like hunger & lust also grew more complicated.

Now, in the blink of a celestial eye, 700 million years have passed (cue Terence Malick's *"Tree of Life"*). Here, humans have gotten the long end of the stick. Their brains are *awesome.* Those simple implements like hunger/satiation, lust/repulsion & strength/fatigue have morphed into an entire toolbox of fancy gadgets. And those gadgets have a name: *emotions*. The same neurotransmitter-based signals that forced C. elegans to eat all his spinach are now signaling all sorts of crazy & unbelievable things. And they're doing it mile-a-minute. If a roundworm's simple signaling system woke up inside a human brain, it would feel like a previously perpetually-recluse hydrogen atom suddenly transported to the center of a blazing sun.

And according to our theory, *all* of those complex, dynamically-applicable human emotions have evolved from distinct proto-emotions that appeared in earlier vertebrates. To begin with, look closely at the value gain/loss judgements that are at the heart of so many primary emotional pairs. What was the original *object of value*, the one that hunger & satiation managed? *Food*. Hunger. Pain. Value loss. / Satiation. Pleasure. Value gain. *Rodney saved the yummy and felt pleasure—even before* eating *the rescued yummy*. (Interestingly, the *vast majority* of our brain's pleasure-producing serotonin comes from one location: our *stomach*—and the serotonin's commute to the brain is signaled by a specially-designated nerve that connects the two organs. Coincidence? Doubt it.)

In addition, these *other entities* we are always judging, Agents of Value—what was the original *other entity* that early vertebrate brains were most interested in evaluating? *Their mate*. Lust. Affection. Agent of Gain. / Repulsion. Animosity. Agent of Loss. *Rodney was angry at the wanderer for causing the loss of his yummy.*

The emotional "bonding" that is triggered by Agent of Gain judgements (which are involved in many emotions beyond affection—like generosity & magnanimity) is likely aided by the specific use of the hormone/neuromodulator *oxytocin*. Research has shown that this neuromodulator is involved in many "empathetic" (aiding/sharing) [10] or affectionate behaviors (it's sometimes called the "love hormone").[11] And the use of oxytocin by our modern Agent of Gain emotions (to aid in bonding with those dynamically-determined Agents) probably has its roots in that more reflexive proto-emotion *lust*.

As in humans, oxytocin appears to be used by earlier mammals to aid in bonding with mates & offspring, thus its expanded (but similar) use in our modern *Agent-of-Gain-related* emotions seems likely. (And this kind of bonding works in combination with belief-based mechanics like *admiration*—and other predictive patterns/assumptions drawn from accumulated or high-impact experiences—in helping to cognitively define individuals & entities as reliable or "trustworthy.")

Decision-making about all of these resource gains & other entities began getting more complicated when —in the middle of that

700 million year blink—advancing creatures got a cool new (but still primal) neurotransmitter-fueled prediction tool & signaling gadget: *fight or flight*. This little device provided a super-useful survival skill: a method for choosing the most appropriate response to immediate danger. *I can take him! Let's do this!* or *No way, man! Run!* Whaddaya know...*a validity judgement—assessing which one of two predictions is more likely to either achieve a gain or avoid a loss*. Fight. Confidence. Prediction success. | Flight. Anxiety. Prediction failure. *Rodney hesitated before reaching into the fire for his possibly-poopy yummy.*

Keep in mind, exercising *fight or flight* is not the same as identifying a possibly-edible fruit and feeling compelled to eat it. That's simple value gain recognition & signaling. You know *exactly* what to do: eat the fruit. But fight/flight is likely tied to our ancestral validity systems because it involved assessing two possibilities that might be best. *If I fight, I might win & live. If I run, I might get away & live.* You *don't* know exactly what to do, you're weighing your choices—*measuring the validity or likelihood of each prediction*.

Another primitive feeling—one that also seems to be tethered to a modern emotional mechanism—emerged during the heart of that 700 million year blink: strength/fatigue. Is this category a little too imprecise? Probably—inasmuch as it doesn't distinguish between an overall state of fitness & simply feeling rested/unrested. But at its core, strength/fatigue represents a more fundamental, action-specific judgement: *am I able to keep going or must I stop?* This judgement is most vital at times when a creature's survival depends on its ability to squeeze every last bit of life-saving action out of whatever physical resources remain—which can be hindered by things like pain & fatigue (feelings creatures typically experience in these survival-challenged moments).

Once again, nervous systems around the globe went back to that oh-so-reliable tool for a little help in these situations: neurotransmitters. Vertebrates got a gift—*endorphins*, which are released during moments of pain, excitement, exercise & fatigue (and others, like orgasm, but let's stay focused). These endorphins are known to inhibit pain, create feelings associated with pleasure, and to be released in moments when we're trying to squeeze the most out of our resources (injury, exhaustion, sex).

Thus, it seems likely that—as the modern human brain emerged—mechanisms rooted in that primitive strength/fatigue feeling & involving endorphins evolved into what we think of as *willpower:* the attempt to "consciously" bolster one of those aforementioned struggling or difficult (or *extra-resource-requiring*) efforts. Science has, indeed, shown that these endorphins can play a key role when we experience both very open & very guarded states [12] (highly-excited & highly-fearful) generated by our primary (*narratively-based*) emotional pairs —making us more capable of taking effective action in each state. And the roots of this kind of willpower mechanism were probably heavily-intertwined with that validity-based proto-emotion we just described: fight/flight.

To understand why, first consider that the validity judgements necessary to take the most-beneficial dynamic & contextually-based action appear to have actually *preceded* the development of true fight/flight (even though we shamelessly gave fight/flight all the credit on the previous page). This validity-based precursor to fight/flight is something we might think of as a *fight/cower* response.

An example of this in early reptiles: *turtles*. (Humans, of course, didn't evolve from turtles, but reptiles & mammals both emerged from the earliest amniotes. Thus, those first versions of reptile brains likely shared many fundamental mechanisms with those first versions of mammalian brains—and similar basic fight/flight responses are demonstrated by both reptiles & mammals.) When certain turtles dynamically *choose* to respond to a unique new potential threat by either biting or retracting their heads [13] (some are not capable of both) they are making one of those contextual *this-or-that* validity judgements that's the basis of fight/flight.

But, as described, turtles don't typically *flee* —instead, they essentially *cower.* If we consider that, according to our theory, all emotions (proto & modern) are part of an encourage/inhibit pairing, then fight's *encourage* response would naturally be countered by a pure *inhibit* response. This is what cowering represents: *inaction & guarding behavior* in response to a perceived threat instead of active & open behavior (biting).

In order for this creature to overcome its inhibitory cowering response and actually *flee*, they would likely need to begin getting injured while cowering & suffer pain. This is because, among these earlier vertebrates, pain or fatigue were required to generate an endorphin response, which is what ultimately helps them to neurally overcome the inhibitory cowering and actually engage in some life-saving fleeing.

One of the interesting things about fight/flight is that it contradicts that seemingly fundamental action/inaction pairing of emotional responses—fight/flight is actually *action/action*. How did advancing vertebrate brains likely achieve this paradoxical pairing? *Endorphins*. Evolution seems to have sorted out the fact that—if you've already determined a threat is *un-fightable*—it's often better to engage in any necessary fleeing *before* you begin to get that endorphin rush from being pummeled while cowering (especially if you're a *post-turtle* vertebrate whose *cowering effectiveness* isn't enhanced by a shell). Thus, in the development of fight/flight, the brain likely began to repurpose those original endorphin-based strength/fatigue mechanisms & use those neurotransmitters to help counteract that *initial* inhibition response generated by fight/cower.

Basically, this means that creatures with more evolved fight/flight responses would've been the first to generate endorphins based on cognitive analysis of externally-perceived threats (those *this-or-that* validity judgements) as opposed to producing endorphins purely based on internally-detected pain or fatigue stimulus. (This kind of development seems to be one of the primary drivers of evolutionary advancement in vertebrate cognition: the growing integration of neural systems that were originally solely devoted to either external or internal sensory input.)

The neurotransmitter/hormone most-commonly associated with fight/flight is epinephrine (aka *adrenaline*, like our body's version of speed) which is typically released in heavy doses when stressed or otherwise physically-aroused by a situation. But epinephrine doesn't appear to help us to *choose* to act or overcome some inhibitory behavior. Rather, it seems to be released once we've *already chosen* to act or simply upon *encountering* the stressful stimuli—essentially temporarily juicing our whole system, allowing us to perform *whatever* act with greater efficiency, robustness, stamina or effectiveness.

Endorphins, on the other hand, were first designed to be pain-blockers (like our brain's version of opiates) and thus naturally work in *direct opposition* to inhibitory instructions and primal urges. Endorphins aren't just there to provide pleasure that enhances performance, but pleasure that also specifically helps overcome inhibitory or contradicting instructions. Thus, fight/cower likely generates only epinephrine in the creature (making it more effective in fighting or cowering, but still unable to flee) while the more-developed fight/flight response generates both epinephrine *and* endorphins. Similarly, endorphins likely play a key role in some of our "guarded" emotions that can require urgent action, like *anger* (which is probably why it can actually *feel good* to be angry sometimes).

These are the reasons why it's more probable that endorphins and *not* epinephrine are the foundation of our actual *willpower* mechanisms. (In fact, because it's juicing *everything* in the brain, epinephrine can sometimes make it *more* difficult for an urge to be controlled by our endorphin-based willpower.)

Why did the involvement of strength/ fatigue's endorphins with fight/flight's validity judgements end up being so great for humans? Because this evolutionary development connected the release of endorphins to those early cognitive systems that would eventually generate validity-tested, emotion-producing, decision-making narratives. This is how it likely became the root of *willpower*—that attempt to "consciously" bolster a struggling or difficult effort. Once these systems—urge-overcoming endorphin-production & action-enhancing narrative motivations—were tied together, human brains could use these unique neurotransmitters to aid in choosing high-priority & sophisticatedly-arrived-at narrative options over powerful primal or emotional urges.

Unfortunately, endorphins are a fairly new discovery (only dating back to the 1970s) and there is not a wealth of broad research on their effects in different neural circumstances. But there is a small amount of endorphin research that provides an interesting window into their willpower connection: research on endorphins & *sleep*. In one study, it was shown that disrupting endorphin input within the human brain while sleeping had no impact on the sleeping brain; it remained asleep & unperturbed [14]—suggesting that endorphins have no role in the sleeping brain.

In addition, a study on cats showed that the introduction of endorphins to the brain during sleep both inhibited lighter sleeping & entirely prevented deep REM sleep [15]—also suggesting that it is unlikely that mammalian brains are using endorphins while sleeping. Thus, those moments immediately after awakening or moments of semi-sleep (like sleep-walking) are likely brief windows into how our brains might behave *without* the benefit of narrative-action-enhancing, urge-inhibiting endorphins.

As someone with a lifetime of sleep issues (sleep-walking, difficulty sleeping long stretches, etc.) I happen to have a good deal of experience facing the world in either semi-sleeping or barely-awake states—in fact, as I've aged my sleepwalking has been replaced by the odd & disconcerting habit of sleep-*eating*. My experiences in both of these (likely endorphin-deficient) states are fairly common, and one thing seems to be particularly true about all of these experiences: I exhibit a significant decline in my ability to express *willpower* over my urges (like eating half the box of cookies or flying off the handle at the slightest irritation, even though deep in my brain I can hear myself clearly saying *don't eat that* or *calm down*).

This willpower deficiency while barely-awake, as mentioned, is not uncommon. My guess is that many readers of this essay have had similar experiences. And although some of the other primary neurotransmitters like serotonin & dopamine are typically less in evidence during sleep, they are still used in some small fashion or another during the whole process of sleeping and awakening, and their presence in the brain does not actually appear to inhibit sleep in the way endorphins do. Thus, endorphins appear to be one of the only primary neurotransmitters that's entirely absent during these episodes of sleep-induced willpower deficiency, also supporting its candidacy as willpower's main neural advocate.

Whenever a narratively-based cognitive desire (don't eat those cookies, don't get mad about that, *control yourself*) is powerfully contradicted by one of those strong emotional or urge-based impulses, endorphins are released and enlisted in aiding the "preferred" narrative desire. The higher the value you can generate for the preferred choice via your story, the stronger the production of endorphins in support of that narratively-reasoned option.

This is why when guys like Aron Ralston (the dude who was wedged in a rocky crevice & saved his life by cutting off his own arm) finally muster up the willpower to slice away, they do so by thinking of all the people they love and want to return to, convince themselves that they will die otherwise and thus must act to see them again. Ralston even thought about people who didn't *exist* yet—namely, imagining his someday child, who might not *ever* exist if he didn't survive. These are powerful & convincing stories—the kind that help maximize endorphin production & win the battle over the very strong primal urge *not* to cut off your own arm.

Another thing about Ralston, whose endorphin system and story were so amazing that he could cut off his own arm: he was one of those *thrill junkies*. In other words, he seemed to get extra-special and addictive pleasure from the endorphin-enhanced joys of risk-taking behavior and physical exertion. This is evidence that he likely possesses naturally-strong endorphin production or benefit, which is partly what saved his life—that and the aforementioned powerful, convincing & endorphin-maximizing *story* that aided him in winning the battle against *not wanting to cut off one's own arm.*

Which is not to say that *all* individuals with strong natural willpower mechanisms exhibit a *thrill-desire*—many other factors are also at play here, such as our ability to create & maintain those powerful narratives/reasoning that help *trigger* the endorphins. In fact, from our theory's perspective, many of the decision-making conflicts that are considered to be mitigated primarily by "willpower" (such as resisting the urge to cheat on a test) are actually a result of our belief systems working in powerful *combination* with mechanisms like our endorphin-based willpower (a matter that will be explored in Essays 4 & 5).

Ultimately, the particular willpower device that we're identifying here can be described in very specific terms: "willpower" is a neural mechanic that (with the aid of endorphins) encourages humans to *consciously choose to endure* (& helps them to tolerate) predicted & ongoing pain/loss in the service of achieving a longer-term personal or broader societal (& often belief-based) gain. This neural mechanic is cognitively triggered when there is a *strong conflict* between a powerful narratively- or belief-based (consciously-considered) goal and a powerful pain-based/loss-avoidance urge or emotion—like hunger, fear, anxiety or anger. (For example, when the goal of saving your life by cutting off your arm *strongly conflicts* with that fear- & pain-based urge to *not* cut off your arm.)

Keep in mind: even when we use willpower to refuse an easily-available gain like secretly downing an extra piece of cake or swiping an unseen $100 from the register (*or having sex with someone other than your spouse*) that willpower mechanic is still essentially helping us to overcome (& tolerate) the predicted (& ongoing) pain of *not* eating the delicious cake or *not* becoming $100 bucks richer (*or that devilish pain of* not *having sex with someone other than your spouse*).

Additionally, as we noted, the *effectiveness* of this mechanic in helping to achieve or choose the narratively- or belief-based goal is primarily determined by a combination of the strength of the emotional response generated by the narrative and an individual's capacity for endorphin production & benefit. This effectiveness can also be *hindered* by the kind of mental fatigue (aka, diminishing brain resources) that can result from being over-worked, under-rested or stressed-out—which likely makes it harder to maintain the cognitive focus necessary for effective (& willpower-inducing) narrative reasoning. (This mental-fatigue-based willpower hindrance does not, however, *totally disable* our human willpower mechanisms in the way that those *endorphin-deficient* sleeping or semi-sleeping states seem to.)

By viewing willpower in these terms, its connection to that endorphin-based *strength/fatigue* proto-emotion becomes even more clear. When those early creatures attempted to muster their quickly-waning resources in order to take that next *survival-aiding* step away from danger despite extreme fatigue or serious injury—which trigger endorphins—what those creatures were really doing was choosing to endure (& being aided in tolerating) the *pain* that was an inevitable consequence of taking that *oh-so-difficult-but-survival-aiding* next step away from the danger. That *pain* is telling the creature: *Don't move, we're injured!* or *Don't* *move, we're almost out of resources!* But the endorphin-based (& primitively cognitive) response is saying: *We'll worry about that later, because if we don't move RIGHT NOW there probably WON'T BE any "later."*

And so, based on all of this, we can imagine how a complex, endorphin-based willpower system evolved from our ancient strength/fatigue mechanism—thanks to that mechanism's interactions with fight/flight and its eventual connection to those cognitive systems that now govern human choices. *Rodney was so angry with the wanderer that he nearly slugged him—he really wanted to, but knew it was a bad idea. Resisting the urge took all the willpower he could muster.*

~

The next likely leap in vertebrates' emotional evolution reveals a truth that even Darwin had a hard time reconciling: everything isn't *always* & *entirely* about us, the individual. In some cases, it turned out that aiding one's own survival meant aiding *the group's* survival. And aiding the group often meant one specific kind of behavior: sharing resources. It also meant helping out fellow group members in a bind—e.g., helping free a trapped species-mate, which is ultimately a *sharing* or *donating* of personal resources like time, energy & risk-exposure; we're *literally* giving something of ourselves.

This gearing of individual action toward benefitting a larger group by encouraging cooperative behavior (essentially, aiding & sharing behavior) was the beginning of social structures. (Keep in mind that in these social structures, individual actions that benefit the group also provide ultimate benefits for that *individual*, whose own survival is supported by the group.) And recently, researcher Alison Davis Rabosky discovered a rare group of desert-dwelling lizards who present the earliest evolutionary evidence of kin-based social behavior.[16, 17] These lizards work cooperatively to build the tunnel structures in which they live (for multiple generations) & share resources, and this openly cooperative behavior is built around kin-based social structures. In other words, these lizards appear to be among the first to share with other genetically "pre-defined" (by kin) Agents of Gain within their species.

In early mammals, there is actually evidence that this aid/share proto-emotion or instinct was applied species-wide. (To be more accurate & less positively-biased, this proto-emotion is better described via its root encourage/inhibit pair: Share/Hoard.) Peggy Mason at the University of Chicago demonstrated that rats will help free a trapped (& unrelated) rat, and also share the yummy chocolate chips that both rats have access to (they will even free the trapped rat first, despite having open access to the yummy chocolate chips).[18] Other recent research in rats has shown that they appear to use mirror neurons to empathically reflect/experience stress & pain observed in other rats.[19] Thus, it seems likely that the mirror-neuron-perceived distress in other trapped (or otherwise stressed) rats helps to trigger sharing's resource-donating twin: "aiding" behavior (a string of dynamic responses that continue to be tested & revised until no stress is empathically perceived in the other rat). All of this essentially represents indiscriminate "altruistic" behavior in which donating resources to any fellow species-member represents an overall survival benefit.

In later mammals (like pack animals & primates), this aid/share instinct mostly grew more discriminatory again—applied only to kin (like those lizards) or other members of tightly-knit social groups, thus allowing more intra-species competition for resources. With the exception of bonobos, who—as proven by Brian Hare at Duke—actually *prefer* to share with *strangers*.[20] It seems that ever-social bonobos value expanding their social circle above all else.

Hare's most recent (& brilliant) bonobo experiments also demonstrated something else very revealing: the sharing-inclined bonobos would *not* share with the stranger if an *actual* food loss wasn't *counterbalanced* by the gain of *actual* social contact.[21] I believe this loss/gain "counterbalancing" is still an essential element in modern human sharing; no matter how powerfully our relationships or beliefs may compel us to share, there is almost always some "maximized" level of loss that will inhibit that powerful urge to share or aid. (The *most-maximized* level of loss is, of course, losing *our lives*—which we're typically only willing to donate in the service of our most dear causes or in aiding our most profoundly-bonded Agents of Gain.)

Even though this kind of kin- or pack-based (or stranger-based) sharing was a more discriminatory application of this proto-emotion in mammals, it was still applied primarily according to genetically "pre-defined" Agent of Gain criteria. And sharing stayed that way (pre-defined) for a long time—until hominins (or likely until some of their closest primate relatives). Thanks to those newly-evolved, awesomely-modular & flexible neural systems, humans added a new trick to our judgements of other individuals/entities: that dynamic tagging of Agents of Gain or Loss (the descendant of Lust/Repulsion).

In other words, no matter *who* you are, if you help me or hurt me I'm going to *remember* that and tag you as a helpful or harmful entity for future reference. This individualistic, dynamic tagging of Agents of Gain also meant that our aiding/sharing behavior could be applied in a newly dynamic way—allowing humans to feel those modern emotions like *generosity* or *selfishness* toward entities that we have specifically categorized as helpful or harmful. Share. Generosity. Donate resources. / Hoard. Selfishness. Protect resources. *The wanderer promised to help Rodney hunt in the morning, and the offer made Rodney feel better about sharing his yummy with the wanderer.*

(Earlier mammals—like dogs—can also remember entity-related gain events that can *ultimately* affect future sharing behavior with that entity & help them make emotional judgements like anger & affection. But I believe those initial gain-providing interactions can actually allow that entity to obtain "pack member" status. And that pack status is still the "pre-defined" neural judgement that determines specific sharing behavior, which is a neural judgement that's different from anger & affection.)

It's important to understand, however, that a human's decision to *share* or *hoard* isn't all about our dynamic Agent of Gain/Loss tagging—because humans have those other powerful behavioral calibrators: *beliefs*. In other words, we can have very specific & hierarchically-organized learned beliefs regarding sharing and apply those in combination with our more primal (but sophisticatedly-dynamic) Agent of Gain or Loss judgements when making decisions about sharing resources or providing aid.

Ironically, from Narrative Complexity's perspective, our human empathy mechanisms (which I do not believe are actually much different from the empathy mechanisms of other primates) only play a tertiary role in human aiding/sharing behavior—behind the roles of those beliefs & Agent of Gain/Loss mechanisms. Consider that "empathy" is ultimately the result of mirror neuron-based systems that reflect visually-perceived "other entity" physical movement & facial expressions from our parietal lobe to our pre-motor & somatosensory cortexes—which allows us to internally experience & interpret those "other entity" physical movements & feelings.

Thus, this empathy mechanic really only provides humans with the *capacity to better judge* (& feel for ourselves) how others are feeling, but empathy does not actually have much impact on how we choose to *respond* to that experience or judgement of their feelings. Our actual response to empathically-based emotions & judgements is mostly determined by cognitive mechanisms like those Agent of Gain/Loss mechanics & belief systems.

Someone says something mean to you. You respond by saying something even meaner & it almost makes them cry. You visually & empathically

—via mirror neurons—identify their sadness & reflexively, to some degree, feel their sadness. Do you apologize or walk away satisfied?

In both cases, you *empathically* perceived their pain—and this empathic perception might automatically trigger at least an *echo* of that primal "aiding" urge—but ultimately, your full behavioral *response* to that perception (& that echo) depends upon your beliefs about concepts like forgiveness, and judgements like whether or not you perceive them as a potential Agent of Loss or Gain.

Returning to the trail of emotion's evolution, although our earliest social mammalian ancestors did not possess this ability to respond to other group members in such diverse & complicated ways, their primitive-but-ever-advancing social structures did more than just foster basic aiding & sharing. These social structures also helped give value to a new commodity that those animal packs & communal groups allowed: *social status* (the acquisition of which provided myriad survival & reproductive advantages). And wherever there's value to be gained or lost (social or otherwise) emotions are bound to be found. Thus, mammalian brains developed a new proto-emotional mechanism that aided in managing & responding to the gains & losses of this new, valuable *social status* commodity.

In the view of Narrative Complexity, advancing mammals who arranged themselves into more complex (non-purely-kin-based) social groups—e.g., pack animals like wolves—likely used this "social status" behavioral/emotional mechanism to accomplish two tasks that are crucial to forming complex social groups: 1) helping to determine "in-group" & "out-group" judgements of individuals, 2) helping to determine an hierarchical order within the group (aka, determine leaders & followers). And, according to our theory, the proto-emotion that played the key role in those tasks was likely a primitive version of our purely-socially-based modern emotion: *Pride/Embarrassment* (an emotion that, as we noted earlier, *requires* an actual audience—or, at the very least, an *imagined* one). Viewed in its proto-form, we might think of this emotional pair as *Inclusion/Ostracization.*

Whenever a potential or current member of a social group (like a pack of wolves) engages in behavior that harms the group or its pursuit of a goal (*like hoarding food or screwing up your role in a group hunt, allowing the escape of soon-to-be-food*) the social group —usually following the example of the leader—will likely engage in some kind of "disciplinary behavior" toward the offending screw-up. This "disciplinary behavior" is essentially a form of "shaming." And the result of this shaming is that the offender "feels" (at least temporary) *ostracization* from that group.

The behavior that's triggered by this proto-emotion (behavior that's demonstrated, for example, when you scold your otherwise-beloved dog for pooping on the carpet) essentially leads the offending individual to "self-ostracize" or engage in behavior that distances themselves (physically and/or socially) from that group.

Conversely, when a potential or current group member engages in behavior that specifically aids the group or its pursuit of a goal (*like impressively taking down the big & elusive target of a group hunt or wisely leading a group of foragers to the perfect location for abundant foraging*) that individual is likely to receive a positive response from other group members (essentially a form of "praise" combined with primitive expressions of gratitude). The result of this positive social response is that the individual experiences a powerful "feeling" of *inclusion* within that group. And the behavior that is triggered by this proto-emotion (which can be observed when you effusively praise your dog for a job well done) is something that we might think of as a desire or willingness to "take center stage" (at least temporarily or maybe even *momentarily*).

This kind of primitively prideful behavior essentially signals a stronger engagement with or commitment to the group, which demonstrates to others that individual's capacity to be part of (or take on a greater role within) the group. This *inclusion-spurred* behavior can also trigger within that individual a *desire* to take on a greater role within that group (something that might ultimately lead a powerfully-prideful underling to challenge the reigning *alpha* for group dominance).

The evolutionary-fitness value of this emotional mechanic is that it both helps to sort out the most group-benefitting individuals from the least group-benefitting individuals, and it helps to determine an hierarchy within that group, which is crucial to highly-cooperative behavior like pack-based hunting or group foraging (cooperative behavior that typically requires both a strong, proven, highly-skilled leader and competent, willing & well-disciplined followers). Despite the obvious evolutionary necessity for this distinct proto-emotion's existence in advancing social mammals, it expresses such a fundamental judgement about how we view ourselves (*part of* or *not part of*) that we barely think of this primitive pair as a true set of "feelings" by themselves.

And, in fact, I believe that our difficulty in identifying *Inclusion/Ostracization* as a distinct & separate proto-emotion that can be clearly differentiated from the experience of *Pride/Embarrassment* is because this feeling has actually *barely evolved* from its primitive form into a distinctly modern one. Why have these particular emotional offspring stayed so uniquely close to their parents? According to our hypothesis, it's because a more complex, capable & modern emotional/social tool took on many of the tasks that Pride/ Embarrassment (& its proto parents) originally handled: those ultra-useful & highly-flexible human belief systems.

We'll detail the evolution of our belief systems in a moment (& explain exactly how intricately Pride/Embarrassment are tied to that evolution) but we've already discussed the powerful role that belief-based emotions like admiration/resentment can play in making decisions about following or not following the lead of someone else. Additionally, in modern social groups an individual's in-group or out-group status is powerfully impacted by whether or not that individual has demonstrated or expressed that they share the group's most important & sacred beliefs.

Because our brain's belief systems are so complex & highly-evolved, they are ultimately much more effective & nuanced arbitrators of social groups & group hierarchies than those much simpler Pride/Embarrassment mechanics. In addition, Pride/Embarrassment can be overly-prone to undesirable results like simply allowing the biggest bully—aka, a disproportionately prideful & shaming-prone individual—to take over a group without necessarily demonstrating all of the skills best-suited for leading the group (fueling the political ascendence of rage-filled despots like Adolf Hitler and narcissistic fools like Donald Trump). Thus —having ceded the task of handling the more diverse & robust management of social groups to our beliefs—that nearly-proto-emotion Pride/Embarrassment was never evolutionarily driven to morph into something more distinctly complex. It's like the *Peter Pan* of modern emotions: *it just never really grew up*.

All of which means that—although there's still a distinct parental relationship between the proto & modern versions—the evolved pairings of Inclusion-Pride *(Gain of Social Status)* | Ostracization-Embarrassment *(Loss of Social Status)* still remain uniquely sibling-like (members of nearly the same "emotional generation"). *When Rodney explained to the wanderer that he was the first person in his tribe to discover this unique yummy—and then observed how much this impressed his new companion—Rodney's heart swelled with pride.*

~

Those frequently-aforementioned belief systems finally bring us to what might be the most crucial & pivotal development in the evolution of emotion, one that likely occurred alongside the emergence of social structures in the heart of that 700 million year blink: disease avoidance behavior—

essentially, primitive *disgust*. Early disease avoidance appears to be based on identifying a specific subset of olfactory data within a larger scent pattern. For example, rats could detect & identify a subset of disease-indicating olfactory data within the larger scent pattern of another rat, which triggered survival-aiding *avoidance behavior*.

(The unique neural mechanics & roots of primitive disgust are well-explored by Hanah Chapman & Adam Anderson in their 2012 paper "Understanding Disgust." Additionally, as their paper notes, humans' & other animals' *distaste* response—primarily spurred by specific stimuli like bitterness, and intended to identify toxicity as opposed to a possible disease-source—is much more primitive & less sophisticated than disgust. [22])

In the view of Narrative Complexity, this neural mechanic—applying a specific, but broadly-applicable subset of data to larger data patterns in order to determine avoidance behavior—is what unites *all* forms of disgust. This mechanic is demonstrated by advancing mammals' capacity to specifically judge, for example, disgust-producing (& possibly-illness-causing) *rottenness* across a wide variety of unlike fruits & meats.

As mammals evolved, different species developed different levels of disease avoidance behavior—likely based on the species' specific natural disease-resistance. (Thus, species with greater natural disease-resistance, like dogs, would require less powerful & broadly-applied primitive disgust responses.) Hominins not only inherited this olfactory-based, disease-avoiding disgust, but they also seemed to possess a particularly powerful version of it—demonstrated in our strong, primal aversion to the scents & tastes of harmful resources like rotten food or feces (stimuli that don't seem to particularly bother the olfactory systems of mammals like those aforementioned dogs).

And since we've mentioned dogs, it seems fair to note the unique disgust response displayed by their cultural counterparts: *cats*. Felines appear to express this avoidance behavior by reflexively attempting to *bury* or *conceal* the offending material (and they even seem to reflexively seek out a burying-favorable location—a pile of sandy dirt or a litterbox—when depositing their own offending material). Disgust's cross-applied-data-subset mechanic is evident in this behavior too: cats will reflexively display this *paw-reach-&-pull* burying action when encountering a range of different kinds of novel (but powerfully-scented) stimuli. I've seen cats do this in response to items as diverse as ashtrays & coffee puddles—despite the fact that these items' overall, complicated scents are much different from each other & from feces.

For early humans, these flexibly-applicable primitive disgust mechanics were so useful that they eventually made a spectacular & crucial leap: from the olfactory systems to our visual & cognitive systems. What spurred this leap? Narrative Complexity hypothesizes that the key event occurred long after hominin brains had already left all others in the dust, when our human ancestors finally did the deed: *making fire*. This discovery now allowed them to cook their food, which ultimately forced our ancestors to develop & nurture an unprecedented ability: eschewing the primal, *hardwired* desire to eat raw meat in favor of exercising the *learned* behavior to wait & eat the meat *after* it's been cooked (and eating the cooked meat offered a plethora of advantages in areas like digestive efficiency, food storage & general health—i.e., *avoiding* food-borne *disease*).

In his 1999 paper "The Raw and the Stolen," Harvard anthropologist Richard Wrangham hypothesizes that the advent of cooking by early Homo erectus populations played a significant role in the evolution of human social systems. [23] Wrangham theorizes that, initially, cooking was primarily used to take greater advantage of underground storage organs (essentially, root vegetables) during periods of food scarcity. He also hypothesizes that the cooking of meat didn't emerge until *after* the cooking of these root vegetables had already made a significant impact on our evolving human social systems. (Although the earliest environmental evidence of cooking with fire—i.e., hearth-like structures in human-inhabited caves—only dates back to around 1 million years ago [24], Wrangham believes that evidence derived from the Homo erectus fossil record suggests that the cooking of underground storage organs might've actually begun around 1.9 million years ago.)

In the view of Narrative Complexity, despite the powerful impact that cooking root vegetables had on the evolution of human *social* systems, this behavior would not have impacted the evolution of human *cognitive* systems in the same dramatic way that cooking meat would have. Essentially, from our theory's perspective, developing & nurturing a preference for those cooked underground storage organs over the raw versions of the same resources presented less of a *cognitive emotional challenge* than developing a preference for cooked meat vs. raw. This is because the cooking of underground storage organs likely made these less desirable (but in times of scarcity, *necessary*) food resources generally more desirable & palatable (i.e., making their consumption much easier & significantly more pleasurable). In other words—when they were initially presented with the choice between immediately consuming raw storage organs and waiting to consume the new & improved cooked versions during periods of food scarcity—our human ancestors' brains did not have to work very hard to convince themselves (& their communal cohorts) that waiting to eat the cooked version was (for a variety of reasons) highly preferable.

In contrast, raw meat was a food resource that was commonly sought out & consumed by our human ancestors—even during periods of resource abundance. Simply put (although, as Wrangham suggests, those early humans probably didn't consume *large quantities* of raw meat) our ancestors actually *liked* eating raw meat, and chose to do so even when raw meat was not a *last-resort* food resource. Thus, unlike those raw underground storage organs—which were probably viewed as an *eat-it-or-die* food resource—raw meat was a food option that early hominins & their primate ancestors had instinctively *enjoyed & desired* for millions of years whenever the option presented itself.

How does all of this relate to those primitive mammalian disgust mechanisms making that spectacular leap from the olfactory systems to our visual & cognitive systems? Well, for starters, it helps to explain why developing a strong *preference* for cooked meat over raw meat would've required more complicated cognitive gymnastics (like those employed by disgust) than simply choosing to eat (& prefer) cooked underground storage organs instead of the raw versions. (And—as we'll discuss in detail on the next page—evidence of our modern disgust mechanisms' strong ties to meat-eating can be found in modern Homo sapiens innate *disgust* toward most raw meat, which is *not* something that most humans tend to display in response to those raw underground storage organs.)

Thus, when our human ancestors initially began to choose to consume cooked meat over raw, they likely needed to employ some of their more advanced cognitive powers—like their advanced version of *willpower*. In choosing to wait for cooked meat instead of simply eating the perfectly yummy & desirable raw meat, our ancestors were demonstrating the ability to exercise their willpower in the service of a *learned & predicted long-term gain* (not just an in-the-moment, inhibition-overcoming, *get-up-&-run!* self-willed impulse). In addition—because these human ancestors did not yet possess those behavior-calibrating & socially-nurtured belief systems that ultimately emerged from this behavior—the emotional mechanic that these early groups of humans likely used to help socially reinforce the advantageous, new *don't-eat-that-raw-yummy-wait-for-the-cooked* behavior was our original emotional-social tool: Pride/Embarrassment, which enabled the effective *shaming* of non-conformers.

This suddenly-useful ability to develop a preference for cooked over raw meat was so advantageous that it quickly (in evolutionary terms) began to evolve into a hardwired, primal *avoidance* or *rejection of (disgust toward)* that raw meat. And the very close association between that socially-enforced *embarrassment* of eating raw meat & hominins' subsequently-evolving, hardwired, primitive *disgust* toward the raw meat likely accounts for the strong *overlap* between the emotional experiences of *socially-based* Pride/Embarrassment *(in self)*, and primitive disgust's modern *belief-based* descendants: Satisfaction/Guilt (*in self)* & Pride/Disgust *(in other)*.

But there was something even more unique about humans' newly-evolved & hardwired *disgust* toward raw meat: this avoidance behavior was based on detecting & identifying a subset of *visual* data, *not* olfactory data. (Two systems that are—as we'll discuss in the next essay—uniquely isolated within vertebrate brains.)

Consider this: we are often repulsed by the *sight* of particularly bloody or "gory" raw meat, but there is nothing about the *scent* of raw meat that causes a similar repulsion (that's how we can tell by *smell* if raw meat is rotten, because we aren't actually repulsed by the *scent* of raw meat unless it's gone bad). In other words, the thalamocortical loop that is at the heart of our consciousness (& whose cortex-based cognitive systems were originally rooted in ever-growing visual systems) now had use of this *data subset/behavior avoidance* technique: disgust. Consider that no other (or *non-cooking*) species seems to be *disgusted* by the *sight* or "thought" (essentially, the thalamocortical *perception*) of anything in particular. Even our near & dear primate relative, a Chimpanzee, nonchalantly *handles* their feces, even though the scent would likely prevent them from eating it. And it's quite clear that no animal other than humans is disgusted by the sights or textures of raw meat. Indeed, this visually-based application of a disgust response appears to be uniquely human.

Once this mechanic joined humans' thalamocortical cognitive toolbox, it began to do some truly amazing things. How? Let's look one more time at what this unique tool, disgust, really does: it uses a broadly-applicable, but rigidly-defined subset of data to evaluate a wide range of resources and determine which ones to accept or avoid/reject—an ability that was neurally-expanded via our learned capacity to resist a primally-motivated short-term gain (raw meat) in exchange for a longer-term gain (cooked meat).

Doesn't all of that sound an awful lot like *beliefs?* And what's that feeling we have toward someone who has violated one of our beliefs? The same as raw & bloody or rotten meat: *disgust.*[25] Avoidance. Disgust.

Belief violation. / Acceptance. Admiration. Belief compliance. *When Rodney* saw *that the yummy was poopy, he winced—and when he* smelled *the poop, he gagged. Then Rodney looked at the wanderer and shook his head, disgusted by the other man's* violation of a solemn truth*: Don't shit where you eat.*

~

How amazing was this meat-cooking behavior—behavior that allowed the extraordinary evolutionary emergence of beliefs? Consider this: those early humans' closest ancestors had likely been eating raw meat for at least a few *million* years before the advent of cooking. This means that those first instances & traditions of consuming (& encouraging the consumption of) cooked meat would have gone against millions of years of *hardwired* urges & desires.

As simple as it seems to us now, this ability to *significantly* self-delay gratification was a profound leap of logic—a kind that no other earthbound creatures had truly made, a leap that I believe marks the real beginnings of humanity as we know it. I've described this self-delayed gratification as *significant* because: a) fire-building & cooking involved the expenditure of additional resources—time, effort & actual physical resources—at a moment that likely often occurred not long after expending significant resources to *acquire* (hunt & butcher) the meat, and b) for most of those early humans that hard-earned raw meat was *already* perfectly yummy & desirable *exactly the way it was.*

This, of course, begs the question: why would any of those early humans even bother to *try* cooking their meat in the first place? One not-so-far-fetched scenario: a winter-starved human ancestor possesses or discovers a frozen carcass that is "accidentally" cooked in efforts to merely thaw—leading to further meat-cooking experimentation & demonstration of additional benefits. However meat-cooking began, the ability to broadly spread & maintain the practice still required overcoming some powerful cognitive and behavioral obstacles. And these factors help to distinguish our earliest ancestors' meat-cooking behavior from the behavior demonstrated in a very recent experiment that showed chimpanzees were willing to exchange a raw slice of sweet potato for a yummier cooked slice by placing the raw item into a simple device that produced a cooked slice after being shaken briefly—a process designed to mimic basic cooking. [26]

(Some might also point to behavior like *seed-caching* in birds as examples of non-human self-delayed gratification, but in these cases there is no current impulse to overcome, and therefore no gratification being delayed. When the bird caches the seeds, it's likely not very hungry *at that moment*. Thus, the cached resource is viewed as an excess—not as a very currently-desirable item whose value increases if the entity expends resources in order to help "improve" the item while self-delaying that current desire.)

And the powerful belief systems that ultimately emerged from this capacity to *significantly self-delay gratification* played an important role in our species' survival during a critical period of evolution. As the modern human came onto the scene 200,000–100,000 years ago, climate was fluctuating frequently & dramatically. In the regions of Africa where those modern humans lived, this climate instability resulted in environments that switched between lush & arid in mere thousands of years. These evolutionary pressures likely favored the selection & survival of human populations with the strongest ability to understand & dynamically adapt to the ever-changing environment by generationally passing-on these populations' ever-adapting knowledge & practices. Such abilities were based in their brains' complexly-modular, problem-solving, language-based capacities, which also allowed for the evolution of beliefs both within those brains *and* within the now-continuous, ever-sophisticating & emerging *cultures*.

And the human brains & cultures that demonstrated the strongest ability to learn & apply these newly-evolving belief systems would've been inherently better at dynamically adapting to the maddeningly-metamorphosing African landscape (we'll give an example of why in a moment). This process of Darwinian selection favoring the "believers" was likely accelerated significantly during the middle of this 100,000 year window via an event referred to as a "bottleneck" in human evolution. This bottleneck was a short period in which severe, sudden cooling of the planet reduced the human population to near extinction.

The plummeting population led to significant reduction in genetic diversity in our species—and recent analysis of the human genome has shown that *everyone alive today* is a descendant of that small pool of humans that stubbornly (& *ingeniously*) persisted along the South African coast during this bottleneck. One of the most provocative & compelling scenarios depicting this crucial moment in evolution is presented by paleoanthropologist Curtis Marean in his 2010 paper about the coastal adaptations that emerged in this tiny group of remaining humans. [27]

Marean hypothesizes that this prehistoric coastal community consisted of possibly only 600 people, and that the keys to their survival were abilities such as the sophisticated use of fire in tool-building, and exploiting the sea & other coastal resources for their primary survival needs. (Including behavior like *harvesting shellfish*, which was only efficient at the lowest tides—e.g., less life-threatening than scouring tidal pools underwater among the crashing coastal waves.) The tool-building & creative problem-solving skills were probably well-enabled by those modular cognitive systems. But some of the other adaptations—such as planning (& relying on) that harvesting of shellfish during low tides—are the kinds of learned behaviors whose powerful predictions would have required that newly-developed & very specific cognitive tool: a belief system. (In Essay #4, we'll explore in detail just how *uniquely* specific this cognitive system is.)

Consider that understanding tide cycles & correlating the movement of the moon to the harvesting of food is not the same as understanding how to build a tool or a fire, which involve direct causal relationships in their construction. These humans could not have understood *how* the moon makes the water move in the same way that they would've understood that striking two stones made a spark that ignited dry grasses—they could only observe and then come to *believe* that there was a correlation between the water & the moon. In addition, this period provides the earliest evidence of humans using red ocher (our inaugural art supply) in symbolic & ceremonial ways—which is more proof of a sophisticated belief system being present in these humans' brains.

How exactly does this kind of belief (*whenever the moon has this appearance/position, the water will be very, very low the next morning*) correlate to that original *data subset/behavior avoidance* technique that it evolved from? The "data subset" here is *the unique appearance/position of the moon that "causes" the water to be very, very low*—a data subset that is compared to the larger data set represented by the moon's & tide's "overall behavioral pattern" (their full yearly, lunar & daily cycles).

Even if these humans were making this prediction purely according to tide patterns instead of using the moon, this would still be a version of comparing a data subset (low tide periods) to a larger data set (the full tidal cycle). Although, because tides vary in a yearly & lunar pattern in addition to their daily patterns, it was likely actually easier & more reliable to recognize the *lowest-tides pattern subset* by using the moon than it would've been via the tracking of water level patterns alone.

Either way, if these humans *weren't* using *some* form of a belief to guide this behavior, then they would've simply been harvesting shellfish essentially randomly: whenever they noticed that the tides were low enough. This obviously wouldn't be a very reliable method for managing vital resource acquisition, and it doesn't seem to represent the kind of *advantageous* behavior that would be such a great way to survive *the world's greatest winnowing of humans.*

The *emotional* role of a belief like "*whenever the moon looks like this, the water will be very, very low*" is exhibited during actual *behavior* when, for example, more-basic urges or desires come into conflict with that belief in choosing an action. Let's imagine, say, that on the morning of the lowest negative tide (which provides that lunar cycle's only opportunity to harvest the least-accessible & survival-aiding mollusks) our coastal human is very, very tired, and thus chooses to sleep late instead of harvesting mollusks at dawn.

When he puts his head back down on his grass mat & chooses to forego foraging, he might use as his lame excuse something like "I will collect shellfish later." And as he says this to himself, our coastal human likely feels a pang of *guilt*: "I cannot collect shellfish later, I should wake up now." (And this guilt is essentially being *disgusted* by one's own behavior.) Unfortunately for his *now-less-likely-to-be-reproduced* genes, this pang loses out to the pang of his comfy grass mat. This *guilt* is produced by violating his strong belief that "*whenever the moon looks like this, the water will be very, very low the next morning.*" (And he saw the moon look *exactly* that way last night.) In other words, he is making a choice that his brain *believes* will likely lead to an ultimately undesirable result (based on a highly valid & valuable prediction trope built from experience & study).

Our coastal human would therefore likely feel this guilt even if he was only harvesting the food *for himself*—eliminating other possible guilt sources, like failing to contribute to his social group or to fulfill a commitment to others. Consider that even if we are the only ones who will likely suffer the possibly negative consequences of our actions, we're still likely to feel at least *a little* guilt or inner-conflict if those actions represent the violation of a strongly-held belief.

The obvious evolutionary advantage of strong *belief-based emotions* in situations like our coastal human's *inner conflict* is that the most-likely-to-survive brains are those that feel enough *guilt* (& exhibit enough *willpower,* whose *endorphins* are unfortunately in short supply during this sleepy inner-conflict) to actually *get up & forage* instead of succumbing to the primal urge for more sleep (which is, again, a lot like waiting to eat a cooked steak instead of succumbing to the primal urge to eat the yummy bloody steak).

Exploring Marean's coastal scenario shows why human populations with the most evolved cognitive belief systems would've likely owned a key advantage in surviving this bottleneck, and it provides the perfect avenue for this essential human trait to emerge as one of the most powerful & fundamental aspects of modern humanity—because all subsequent human evolution sprang from this harshly-selected tiny population of our best "believers."

Making efficient, reliable predictions about our world based on learned (but not entirely *provable*) correlations between events that often have mysterious, but observable relationships—and the development of a specific cognitive system devoted to this mechanism—is at the root of what separates us from all other animals. Consider that *many* other creatures—birds, aquatic mammals like dolphins & whales, elephants, other primates—have the modular neural capacity for language, and can display the profound behaviors, emotions & even the learned, generationally-fluid *traditions* that can result from such a proto-linguistic capacity (however rudimentary). But they do not have *beliefs*. And I propose that it is our beliefs, and the emotions that they engender, that truly make us *human*.

~

Interestingly, *all* of these aforementioned primitive emotional mechanisms are *still* a part of our emotional kingdom; these original systems remain almost fully intact. In fact, they are still the *rulers* of that kingdom. These proto-emotions (which we now think of as essentially *urges*) are often the last obstacle that any narratively-based decision must confront before action is taken. And the highest level of any urge will almost always supersede *any* narrative desire.

If you are at any of the urge extremes—starving/parched, in the throes of lust, completely exhausted, repulsed by rottenness or in the grip of fight/flight—those primal desires will very likely be prioritized over your narrative goal (unless you've developed—or were born with—a *wicked* willpower mechanism). This dominance of our ancestral urges over their modern offspring offers unexpected proof of an age-old truism: *we'll always be your parents & we'll always know what's best for you.*

Emotion, Meet Modularity

How, then, did we develop our modern, complex emotions from these primitive proto-emotional pairs? Well, that requires some speculation *about* the speculation, but since we're already deep in our "what-if" rabbit hole, let's keep digging...

Our filmstrip slips into one last flashback from that 700 million year blink: the long stretch when early mammal brains were morphing into the human one. This is likely the time when all of those uniquely modular neural structures (discussed in the previous essay) began to evolve. And it was this modularization of basic data & larger "ideas" that lit the fuse that led to our emotional explosion.

Think of it this way—those early mammals were actually pretty smart critters. They could *remember* stuff and make use of it later. *Check out that tiny-brained mouse memorizing the fancy maze that leads to the yummy.* And evolving emotions played a big part in this memory device. The *pleasure* of the yummy helps encode the *pattern* of the maze into the mouse's memory. But those mice-like early mammal minds had a flaw: non-modular data structures—a result of their neural limitations. In the mouse's brain, that memory of the maze isn't a long sequence of linked-but-independently-associative turns, it's one big pattern.

This is why, when Mr. Mouse encounters a similar-but-different maze—e.g., the same exact first half, but different thereafter—the mouse will not likely recognize that the mazes are *partly* the same. He'll either ultimately think of them as entirely different mazes, or exactly the same one (possibly leaving the mouse continually baffled whenever he reaches the different second half—at least until he finally starts thinking of it as *an entirely different maze*).

This means that those pain/pleasure mechanics are still pretty broad in their application—always associating themselves with large, highly-detailed data patterns. But as mammals' neural structures evolved and data became more modular, emotions were able to associate with those modular & more specific pieces of data. These newly diversified associations between feelings & data likely helped emotions to differentiate in purpose & application as they grew more interwoven with specific kinds of data modules. (And as mammals began to employ evermore complex proto-emotions, those emotions' use in encoding specific data with specific "values" might've actually served to *aid* the emergence & evolution of those increasingly-modularized mammalian cognitive systems.)

Thus, using these evolved modular systems, a dog can learn to *symbolically* associate the first step in a sequence with the *actual* pleasure derived from the last step. *Pavlov's dog*: ring the bell and the dog salivates excitedly in anticipation of the predicted food pleasure, *not because he wants to eat the bell.* (For the mouse, seeing & recognizing the entrance to the *previously-cheese-producing* maze makes him interested & engaged, but it likely doesn't give him *pleasure*—the actual pleasure is still reserved for successful navigation & yummy consumption.) Therefore, in those more-evolved mammals like dogs, anticipatory emotions are now possible: fear, excitement, confidence, anxiety. And these symbolic inanimate objects likely also allow for symbolic *entities*: Agents of Value. *Viola!* Anger, gratitude, affection & animosity join the kingdom.

(And based upon my distinct childhood memories of our beloved family guinea pig, *Cupid*—who consistently demonstrated a *Pavlovian &* excited squeal merely upon *hearing* the plastic-crackling of the bag that contained her cherished parsley—it seems that the first examples of this emerging capacity for emotional/neural modularity & rudimentary *symbolism* appeared rather early in mammalian evolution.)

By the time humans arrive in our story, this modularity has gone gonzo. We can do all of this symbolic, predictive & other entity stuff *way, way* better. A massive cerebral cortex allows far more data to be stored for reference, comparison & analysis. Advanced neurons with more connections & more sophisticated associative powers enable data tagging & comparisons to be done with greater precision, and allow our predictions to become vastly more complicated. Emerging research suggests that these modern pattern & prediction mechanisms even involve our ancient cerebellum.

And our dizzyingly complex use of those age-old neurotransmitters—combined with immaculately-tuned areas like our insula, amygdala, orbitofrontal & anterior cingulate cortexes—allow for complex new ways to use those pain & pleasure responses. Guilt, satisfaction, envy, admiration, greed, jealousy, melancholy, all the blends & hues —all are now possible. In addition, those long-evolving *mirror neurons* allow *empathy* to help our minds incorporate emotional data that is physically-expressed by others.

The Mothership has arrived. And she has a passenger: consciousness. Which probably means that Descartes' elegant definition of "being" (after all these centuries) is in need one small edit: I think *and feel*, therefore I am. It's a little less succinct, but maybe a little more true—after all, without *love*, what are we?

A Ghost in the Machine

And so, our filmstrip fades to black, the music swells and...wait a minute—what's that? You *feel* something? You mean the *music* made you *feel something*? Almost forgot about that—music. Pretty cool stuff. And maybe the *coolest* thing about music: we're born with it.

Before you worry that we've suddenly gone wildly off track, *don't*—this is the perfect place to conclude our epidemiological examination of emotions. That's because (and although it's silly, I'll remind you again —we're *speculating* here) music seems to have a very special role in the blueprint of our emotional kingdom. It seems to be a kind of *pattern primer*. Remember that exciting part of the movie *"Contact"* when the crazy-brilliant, recluse mogul sends Jodie Foster the *primer* (a mathematic key or decoder) that allows her to interpret & implement the hopelessly-complicated alien blueprint? In our filmstrip, the human brain's myriad narrative-building, emotion-generating mechanisms are the blueprint—and music is a pattern primer that helps us to interpret & use it.

Music has two vital qualities. One: it is a data pattern that simultaneously accounts for defined "vertical" or parallel relationships between its elements (chords) and defined "linear" or sequential relationships between its elements (melody). Two: the various pattern combinations resulting from these vertical & linear data relationships produce *emotions*. In other words: linear narratives (melody) whose multiple layers can be woven together (chords) to produce emotions.

Thus, music looks like a genetically pre-programmed way for our brain to show itself how to use its "blank slate" narrative & emotional mechanisms (whose pattern-analysis & predictions require *recorded data* to really get rolling). Music is a primer for the blueprint associating patterns with emotions—which is the first thing that our vast, initially-empty data-banks needs to learn in order to begin filling it with that crucial recorded data & learned rules.

One of the main ways in which this musical primer helps to build our systems of cognition is likely through the application & interpretation of *inflection* in spoken language (a matter explored in Essay #4). Inflection (which is essentially founded upon those inborn musical rules) allows infants & toddlers to associate emotional values with verbal utterances *before* they've developed a true capacity for language—thus helping to construct that initial basic syntactic framework necessary for developing the complex (& primarily *learned)* linguistic & cognitive processes that sustain human consciousness.

Music is a ghost in the machine. Because our DNA can't pass along the actual data that human brains use to create all that magic, it instead sneaks into the operating system all the pre-programmed emotional responses to the patterns of music. And this pattern primer likely helps our developing brains to make those all-important associations between the mechanisms that analyze complex patterns & predictions (narratives) and those mechanisms that produce behavior-guiding emotions.

From this perspective, it appears that the tools of music might actually help to "jumpstart" (or at least "lubricate") the observe-analyze-respond loop that is the engine of our consciousness. Music, however, obviously isn't the only primer available to us (deaf humans' brains seem to get started up just fine without it). Conveniently, DNA is a pretty spectacular courier of information. It's easy to imagine lots of visual, tactile, olfactory, etc. pattern primers (e.g., those specific emotionally-correlated *facial expressions*) being packaged in our genes in order to help young minds usefully associate emotion with experience—ensuring plenty of redundancy for a resource of vital importance.

Nonetheless—whether or not it's merely a blind spot darkened by a false belief—music *seems* to be uniquely capable in its role as our gateway drug to the addictive & ceaseless pleasures (& pains) that come from associating patterns with emotion. Which is why it feels so…*lifelike*. Why it's so extraordinarily powerful in imprinting a specific moment with its specific feelings—which can still be distinctly reproduced when the music is heard again, even a lifetime later. Music doesn't just know *how to work* the system, it helped *build* the system.

That's also likely why music feels so fundamentally *symbolic* to us, why it so often seems to express how we feel better than we can actually express with words. Words are *almost* perfect. Music is *sublime*. And of course it is. It's some of the most-ancient, most-eloquent code in the universe—light years before the code of words.

And when these different emotion-producing tools—the words & syntax of our internal narratives and the patterns of music—are working synchronously *together*, some magical moments can occur. This is likely why we tend to seek out music that's *mood-appropriate*. From experience, it seems quite clear that there is a uniquely *interactive & amplified* emotional effect when we listen to music whose emotional equations/patterns match the emotional equations/patterns of our internal narratives (basically, when we listen to music that "expresses" how we feel about or *want to* feel about our lives at that moment).

Words are, indeed, more versatile & programmable—wizards of the high-speed modern, modular brain. But just as modern emotion's ancestors (urges) still speak to us most clearly, music knows us in a way that words do not. When our minds, at last, are nearly-gone of all those magnificent associations & cross-associations of data devoured in our lifetime—one set of associations typically remains beyond all others: the musical ones. Thought leaves us, but song often stays—nearly to the end.

And if you believed that all of this complex neural magnificence was bestowed upon us by some vast & unknowable intelligence—as you might suspect, *I do not*—but if you did, then you might assume this *musical persistence* was its parting gift to our consciousness. That before our consciousness goes, before it fails—as it must—it still retains something ancient & sublime, something that might allow us to remain in some way *human* until the end. For music mimics human life at its most fundamental: the association of data, *experience*, with emotion.

Yes, in the end, we are merely the *courier* of a *smaller courier* (that brilliant DNA). But what gives our experience—our journey delivering this valuable parcel to the next generation—what gives that journey any meaning to us at all is the emotion we feel along the way. Does it matter that the ultimate purpose of these emotions is simply to make us a better courier, and *not* actually to imbue our journey with meaning? I don't know. *Does it?* Does it matter to *you*? Now that you are contemplating these possible truths—do you love your mother less? Is there no more anger when you think of that President whom you hate? Is there nothing you desire any longer? Emotions *are* confirmation bias: they matter to us because they *feel* like they do. Thus, the gains & losses, Agents of Value, and validity that our emotions paint our world with—and the beliefs they reinforce—they all matter too, *because it feels like they do*.

And so it is. We are a paradox of emotion—feeling like our lives matter because we feel like our lives matter. Well then, fuck it: *feel*. And let the *logic* of your *emotions* lead you. Let them make you believe that everything in this life that you *feel* like you believe actually *matters*. Find the love. Go after happiness. Why not? If you're stuck inside of a finite and ultimately inescapable & indecipherable illusion, only a fool would hope that illusion becomes a nightmare.

We're here, my fellow humans. *We're in it*. And we're only in it once. We might as well dream the dream.

###

Very Complex Emotions (Mixes of Primary/Complex Emotions)

Disappointment / **Delight**
[*Surprise* + *Pain*] / [*Surprise* + *Pleasure*]

Frustration / **Amazement**
[*Surprise* + *Pain* + *Anger*] / [*Surprise* + *Pleasure* + *Gratitude*]

Horror / **Awe**
[*Surprise* + *Disgust/Disdain*] / [*Surprise* + *Pride/Admiration*]

Despair / **Hope**
[*Confidence* + *Fear* + *Guilt*] / [*Confidence* + *Excitement* + *Satisfaction*]

Melancholy / **Joy**
[*Pleasure* + *Sadness*] / [*Pleasure* + *Happiness*]

Jealousy / **Covetousness**
[*Disgust*-With Other Entity+ *Pain* (Gain: Other Entity; Loss: Self)] /
[*Pride*-In Other Entity+ *Pain* (Gain: Other Entity; Loss: Self)]

Resentment / **Envy**
[*Disdain*-For Other Entity+ *Pain* (Gain: Other Entity; Loss: Self)] /
[*Admiration*-Of Other Entity+ *Pain* (Gain: Other Entity; Loss: Self)]

* This list does not represent a *complete* accounting of all the various mixes of Primary/ Complex Emotions. There are ultimately a *wide array of different emotional states* that can result from various combinations of & intensity levels within our Primary Emotions.

Endnotes:

p. 25:
1. Kahneman, Daniel, and Amos Tversky. "Prospect theory: An analysis of decision under risk." *Econometrica: Journal of the Econometric Society* (1979): 263-291.

2. Tversky, Amos, and Daniel Kahneman. "Advances in prospect theory: Cumulative representation of uncertainty." *Journal of Risk and uncertainty* 5.4 (1992): 297-323.

p. 26:
3. Rolls, Edmund T. *Emotion and decision making explained.* Oxford University Press, 2014.

p. 27:
4. Smith, Craig A., and Richard S. Lazarus. "Appraisal components, core relational themes, and the emotions." *Cognition & Emotion* 7.3-4 (1993): 233-269.

5. Plutchik, Robert Ed, and Hope R. Conte. Circumplex models of personality and emotions. American Psychological Association, 1997.

p. 31:
6. Valdesolo, Piercarlo, and David DeSteno. "Synchrony and the social tuning of compassion." *Emotion* 11.2 (2011): 262.

p. 43:
7. Farb, Norman AS, Hanah A. Chapman, and Adam K. Anderson. "Emotions: Form follows function." *Current opinion in neurobiology* 23.3 (2013): 393-398.

p. 56:
8. Daniel Togo Omura, "*C. elegans* integrates food, stress, and hunger signals to coordinate motor activity." *Massachusetts Institute of Technology*, June 2008.

9. David E. Nichols, Charles D. Nichols. "Serotonin Receptors." *Chem. Rev*. 2008, 108, 1614–1641.

p.57:
10. Sarina M. Rodrigues, Laura R. Saslow, Natalia Garcia, Oliver P. John, and Dacher Keltner "Oxytocin receptor genetic variation relates to empathy and stress reactivity in humans." *Psychological and Cognitive Science*, Proc Natl Acad Sci U S A, 2009 December 15; 106(50): 21437–21441.

11. Jorge A. Barrazaa, Paul J. Zak. "Empathy toward Strangers Triggers Oxytocin Release and Subsequent Generosity" *Values, Empathy, and Fairness across Social Barriers: Ann. N.Y. Acad. Sci*. 1167: 182–189 (2009).

p.59:
12. Gray, Jeffrey A. "Brain systems that mediate both emotion and cognition." *Cognition & Emotion* 4.3 (1990): 269-288.

13. A. Herrel, J. C. O'Reilly, A. M. Richmond. "Evolution of bite performance in turtles" *J. Evol. Biol. 15* (2002): 1083–1094.

p.61:
14. Netz J, Medert HA, Arndt JO. "The opiate antagonist naloxone does not arouse man from natural delta sleep." *Psychopharmacology*, 1986; 90(2):263-7.

15. King C, Masserano JM, Codd E, Byrne WL. "Effects of beta-endorphin and morphine on the sleep-wakefulness behavior of cats." *Sleep*, 1981 Sep;4(3):259-62.

p.64:
16. Rabosky, ARD, A. Corl, Y. Surget-Groba, H. Liwang, and B. Sinervo. 2012. Direct fitness correlates and thermal consequences of facultative aggregation in a desert lizard. PLoS One 7: 1-8.

17. Davis A.R., A. Corl, Y. Surget-Groba, and B. Sinervo. 2011. Convergent evolution of kin-based sociality in a lizard. Proceedings of the Royal Society – B. 278: 1507–1514.

p.65:
18. Bartal, Inbal Ben-Ami, Jean Decety, and Peggy Mason. "Empathy and pro-social behavior in rats." *Science* 334.6061 (2011): 1427-1430.

19. Atsak P, Orre M, Bakker P, Cerliani L, Roozendaal B, et al. (2011) Experience Modulates Vicarious Freezing in Rats: A Model for Empathy. PLoS ONE 6(7): e21855. doi:10.1371/journal.pone.0021855.

20. Hare, B. Kwetuenda, S. 2010. Bonobos voluntarily share their own food with others. Current Biology. 20, R230-231.

21. Tan J, Hare B (2013) Bonobos Share with Strangers. PLoS ONE 8(1): e51922. doi:10.1371/journal.pone.0051922

p.70:
22. Chapman, H. A. and Anderson, A. K. (2012), Understanding disgust. Annals of the New York Academy of Sciences, 1251: 62–76. doi: 10.1111/j.1749-6632.2011.06369.

p.71:
23. Wrangham, Richard W., et al. "The raw and the stolen." *Current anthropology*40.5 (1999): 567-594.

24. Berna, Francesco, et al. "Microstratigraphic evidence of in situ fire in the Acheulean strata of Wonderwerk Cave, Northern Cape province, South Africa."*Proceedings of the National Academy of Sciences* 109.20 (2012): E1215-E1220.

p.74:
25. Chapman, Hanah A., and Adam K. Anderson. "Things rank and gross in nature: A review and synthesis of moral disgust." *Psychological bulletin* 139.2 (2013): 300.

26. Warneken, Felix, and Alexandra G. Rosati. "Cognitive capacities for cooking in chimpanzees." *Proceedings of the Royal Society of London B: Biological Sciences* 282.1809 (2015): 20150229.

p.75:
27. Marean, Curtis W. "Pinnacle Point Cave 13B (Western Cape Province, South Africa) in context: the Cape floral kingdom, shellfish, and modern human origins." *Journal of Human Evolution* 59.3 (2010): 425-443.

The Night Shift

Out Of The Darkness

It is a doppelgänger of our waking life—the universe of our dreams. When we're immersed in them, those dreams can *feel* as real as the world in which you're reading this right now. This is because, in both places, one thing is the same: we sense that we are conscious. And we are—but they are very *different forms* of consciousness. And there are reasons for this, because in our brain, there are reasons for *everything*.

So, why do we dream? In the view of Narrative Complexity, that's an answer that the mechanisms themselves reveal (particularly when viewed alongside the work of dream research pioneers like psychiatrist Allan Hobson [1, 2]). One of the beauties of the brain's elegant systems is that once you've laid out their basic mechanisms, some of the more specific aspects of consciousness—like dreaming—start coming into focus. The sharpest view of our dream world: it's the *night shift* of our consciousness.

What is this night shift's job? During the day shift, our consciousness' primary purpose is to guide our decisions & actions. When we're asleep, of course, we're not *acting* on anything. But determining actions *in the moment* isn't the only purpose of our consciousness. It's also using data derived from daily experiences to help prepare its systems for *future* actions. As it processes this incoming data, our consciousness is simultaneously shaping both how these systems will respond in the future and the scope of those responses—helping establish or reinforce cognitive routines and connections within stored data via application and association of that incoming day-shift data (*topics of Essay#4*).

According to our hypothesis, the night shift (aka, *dream consciousness*) specializes in that last part: *applying* & *associating* data in order to help shape both how these systems will respond in the future and the scope of those responses. But when we're dreaming, there's none of that incoming *real-world* data for our consciousness to process. Therefore, instead of applying & associating incoming (via sensory organs) real-world data, our dream-state seems to primarily process & associate *stored* data (and the *subsequent* data produced by the stream of dream consciousness that's generated and sustained by that stored data.)

When that first piece of stored data (likely something recent or high priority) slips into the mechanisms of our *internal dialogue loop* during our dream state—REM sleep—the associations & narrative-building that induce dream consciousness begin. Why is our brain engaging in this (highly fluid) generative & associative process while we sleep? We'll be delving into those answers in greater context throughout this essay, but the bottom line: consciousness is a prediction/problem-solving machine, and its mechanisms are more flexible & usefully-applicable if they are conditioned to allow for a broader, more diverse range of possible predictions or solutions. (This kind of purpose is also supported by neuroscientist Eric Hoel's insightful 2021 paper "The overfitted brain: Dreams evolved to assist generalization." [3] Hoel's framing of the matter as an "overfitting" problem spurred me to revisit this essay and make clear here that the *applying*—essentially, the narrative element—is just as beneficial as the *associating* in our dreams.)

Here's Where Things Get Weird

Once that first piece of data has started the loop during dream consciousness, literally anything is possible. From our systematic view, their are likely two main reasons for this wildly-open dream world. One, our narrative loop is using the same associative subconscious mechanic that it always uses when processing data. This means that even though the most recent or currently high-priority data is most-likely to both begin & continue to reappear in our dream narratives, *all data* is up for grabs through that magical network of language, emotion & memory. So if your dream suddenly seemingly veers off track, what's really happened is that your brain has merely made one of those miraculous leaps of association into another buried network of ideas, memories & images.

As we said, there are likely *two* (at least) reasons for this wildly-open dream world. The second reason is at the root of why the first reason (our associative neural network) results in more dream-randomness than awake-randomness. That's because, in our dream state, there's none of that *actual* environmental data to support our narrative loop. Our physical environment is what keeps us literally & figuratively grounded; it's the primary factor in the consistency of our wakeful consciousness.

When we're awake (& not daydreaming) our consciousness is framed by what we perceive & identify to be most important/relevant/useful in our environment. Think of this environmental data as *the setting* of our narrative loop. Most wakefully-conscious narrative events happen in the context or in relationship to the context of our immediate physical setting. In our dreams, there *is* no incoming data depicting a physical environmental setting. But our narrative loop still requires one. In a way, it more than *requires* one, our narrative loop simply *assumes* there's one (just as it's the fundamental nature of this loop to *assume* there's some kind of *story* to be derived from or found in this setting). And just like so many other elements in our narrative-building process—if our brain can't find the appropriate data to complete the picture, it makes it up.

(Re: that external sensory input, as noted by our chart in the Appendices, *Distinct States of Consciousness & Non-Consciousness*—the only external input system that appears to be almost *entirely* shutdown during R.E.M. & non-conscious sleeping states is our external *visual* input. External olfactory, auditory & tactile input systems are all obviously repressed during sleeping states, but still remain partly-active and are capable of processing & responding to intense external stimuli—which is why data from external stimuli like loud noises, powerful scents, and *sudden unexpected wetness* might find its way into your dream just before it helps to spark you awake.)

So, in the absence of *actual* environmental data, our dreams get their own brain-painted settings (& people, which are really just a unique part of the setting). And the brain doesn't have to go very far to get the necessary data—all the stored data used in dreams has lots of attached visual, auditory, etc. information, which the brain can make use of when building a scene. Usually most of this attached sensory data is either ignored or discarded in the process of creating our waking conscious narrative—overridden by the more *neurally prestigious* actual incoming sensory data. However, not all of this already-stored sensory data is discarded while running our waking narrative loop. Whenever, for example, we have the normal gaps in perception during our observation of the "real" world, but suddenly require that unperceived data to complete a current or retrospective conscious rendering of the moment, we use appropriate pieces of that already-stored sensory data to fill in the gaps.

This is why you might claim—upon almost immediate recollection—that *the man who stole the woman's purse* was wearing a *red cap*. When in fact, your brain ignored the color in the actual moment, using the more generic *baseball cap* visual contours and/or linguistic tag when it perceived & recorded the event. Thus, upon that nearly-instantaneous recollection, a common red cap from your memory is used when you replay the moment in your mind and return to that element in detail. Until you saw the woman's purse being stolen, your brain didn't care what color the man's cap was, and this data's low priority led your mind to neglect recording it. And after the purse was stolen, in those few seconds it took for the brain to re-prioritize & analyze the details of the man's appearance, he was lost in the crowd. This vague version of the man is simply not good enough for your brain when someone immediately asks you: *What did he look like?* Our brain knows that in situations like this the questioner—and thus *we*—want a better answer than *A guy in a jacket and cap*. And, as explained earlier, our brain doesn't need to try very hard to give these demanding audiences a better answer. To make matters even more dodgy, once the brain gives this "enhanced" answer, it has a tendency—because of the mechanisms of memory—to make a quick believer of even *itself*. If you think this means that we typically trust our memory far too much, well—I don't think I could argue with that. (We'll discuss these kinds of memory mechanics in much greater detail in our next essay.)

And Let The Weirdness Be Unfettered

Getting back to our *night shift*...when we're dreaming, this "fill-in-the-blank" mode of setting & entity construction becomes the sole engine for the creation of our perceived environment. And because there's no actual environmental data to override this mode, our dreams are wildly fluid in their settings & contents. As soon as the narrative loop makes a big leap to another neural network, the dream's elements (including *people's identities*) can be morphed by whatever data is most closely related to the newly-pinged network.

And when we're in this dream mode, our brain doesn't seem terribly concerned with making believable situations out of this random data and typically doesn't even attempt to make sensible narrative transitions between the events & settings. Although our dreams often seem to employ their own habitual & vague (i.e., incredibly *lazy*) "solutions" to these nonsensical leaps. For example, many of my own dreams take place in massive-but-ambiguous "hotels" or "complexes" or "campuses" within which all of the variously incongruent circumstances/settings are seemingly, fuzzily & nonsensically contained.

This *unfettered weirdness* or unbelievability in our dream scenarios is likely because of another apparent quirk of our dream-state consciousness: the normal path of our narrative loop seems slightly altered, entirely skipping (or not "awakening") those *validity* mechanisms in our brain. (Mechanisms governing judgements about likelihood & plausibility, and producing emotions like anxiety & doubt—as discussed in the previous essay.)

This altered loop actually makes sense when we look more deeply into those real purposes of our dreams. As we mentioned, when we're awake, our consciousness' primary purpose is to make quick, reliable decisions, then move on to the next one. This means that there's lots of possibly valuable information that might not get an extended pass through our associative processes as we plow through the day. It also means—during those often *redundant* days—that incoming data is repeatedly running through similar grooves in your cognitive routine.[4] Both of these issues can ultimately contribute to the same result: narrowing the scope of possible future predictions & solutions generated when applying & associating incoming data.

At some point in our evolution (or, more likely, in the evolution of a few stages of mammals before us) the algorithms of time & DNA discovered that—since this consciousness system wasn't doing anything while we slept, and because we could run it without expending the energy to create (and risk) actual resulting actions —this big window of non-wakefulness was a great time to conduct that diverse, broadly-based *applying* & *associating* of stored data. And so, the brain put *the night shift* to work.

This night shift's main job is to generate as many unlikely-but-possibly-useful applications of & associations with recent or high priority data (and the random data that flows from it). The result is that when you wake up, there are new (or stronger existing) associative pathways & broader (or *less-rutted*) cognitive branches (aka, *rule sets*—discussed at length next essay) thanks to those dream-processed associations & applications.

Thus, because dreams are essentially meant to help us broaden the scope of possibly-useful data applications & associations, inhibiting or switching off our validity mechanisms aids this by unleashing an uninhibited range of narrative possibilities and an unlikely parade of data associations —which are especially fluid in the absence of a real-world physical environment. When we're conscious, our validity mechanisms help us to reject & limit (or *be very nervous about trusting*) narratives that don't seem plausible, because they're not typically useful when making decisions. If we did this in our dreams, the inconsistencies & hyper-fluidity in plot & setting that are inherent to dreams (and useful for broad-ranging data application & association) might lead to a near-constant "rejection" of our dream narratives.

In essence, our narrative loop would create the kind of "insanity" apparent in individuals like schizophrenics, whose neurological disfunction produces myriad narrative inconsistencies—resulting in an insanity rooted in the anxiety created by an ongoing negative-validity judgement of the world around them. When we shut down or avoid our validity station while dreaming, we are free to generate & experience an illogical, invalid, but fluidly-associative world without the anxiety that our brain is trained to produce in the face of such a world.

Although it *seems* like our dreams are often *filled* with anxiety—in the view of Narrative Complexity's emotions hypothesis, the emotion we are actually feeling is *fear.* And this emotion (as discussed in the previous essay) is very closely-related to anxiety, but fear judges only the potential value loss, *not* the *validity* of that value-loss prediction. Thus, our truant validity in our dreams likely intensifies the fear & excitement—because our dream-view of those potential gains & losses is unmitigated by any judgement of their likelihood. When it seems like we're constantly *totally stressed* in our dreams, it's because every potential loss we imagine in our unawake narratives feels, essentially, inevitable. (And that gauzy *bliss* we feel in our happiest dreams is likely a result of perceiving their glorious gains as, essentially, certain & perpetual.)

If you're looking for *more* evidence that those validity judgements have gone AWOL during our dreams, consider one of the primary & most-recognizable of our validity-based emotional responses—the response to a highly unlikely (or unpredicted) result: *surprise*. In our dreams, *nothing* seems to surprise us. A talking, human-sized elephant wearing your mother's raincoat and begging you not to eat the birthday cake? *Sure, we'll buy that.* Sights & scenarios that would set-off full-shock validity alarms during the day are totally *unsurprising* during the night shift.

Then Let The Weirdness Be Forgotten

Another one of the other likely side-effects of these absent validity mechanisms: our dreams tend to be awfully slippery critters in terms of memory. When we record a waking memory, the validity of the narrative is a key player in how this data is remembered. The less valid the data, the less likely it is to get the call back when related data enters the system later. Therefore, an awful lot of our dream-produced narratives are basically labeled with the *don't bother* tag in our memory, if they were recorded at all.

Which is okay for our brains, because even if the actual narrative has been kept from full recall, the new associations & cognitive impacts—no matter how faint—still remain. And those results are all the brain was looking for, to add some new possible pathways for our next waking attempt to generate a useful or unique prediction or solution. Plus, y'know, it *didn't* actually happen; it's *not* valid data and *doesn't* provide real evidence about the world—so we're much better off *not* remembering it as real. This is probably another reason why evolution found it useful to shut down these validity systems while dreaming: if they were active, a dream that accidentally seems very real and passes the validity test might ultimately (& incorrectly) be judged later as an actual event. However, this absence of a validity tag doesn't mean that *none* of the dream narratives are well-remembered. The value judgement & tagging systems that are still up-and-running in our dreams are another key player in how strongly a memory is imprinted, so dreams that involve high value events (which create strong in-dream emotions, e.g. nightmares) can still result in narratives that can be easily recalled & replayed in our minds when awake.

This in-dream combination of "online" *value-based* emotional/memory systems & "offline" *validity-based* emotional/memory systems is likely why we can *remember* some dreams & yet still intuitively *distinguish* those dreams from some other *real* memory—because those dream memories lack that neural validity, which allows us to consciously "perceive" their *falseness*. Of course, if you start remembering a particularly-believable dream while awake, and begin too-closely associating it with some *actual* memories while those validity systems are now "online"—it can become possible over time (& the course of a very hectic life) to start accidentally tagging an old dream narrative with the tiniest bit of *wrongly-placed* neural validity. And once this dream has its toe in the door of *truth*, it can be a slippery slope to becoming a kind of *truly remembered* reality.

Returning to those more *useful* powers of our dreams—due to another key player in our memory (*repeated recall*, whose effects will be discussed at length in the next essay) *any* processing or pinging of that recently-consumed (*real-life*) data during our dreams likely aids in that data being more well-remembered and easier to recall. Which might explain why it's been shown that studying immediately before sleeping aids in remembering & applying that specific material later. By studying the material immediately before sleeping, you're helping to ensure that this very-recent data will be used during the night shift, and thus, be easier to recall and use the next day.

About Those Wondrous Walnuts…

When we talk about these individual mechanisms of consciousness (like validity) being regulated—in essence, the regulation of those systems that both *determine whether we are awake or asleep*, and *how our consciousness-producing internal dialogue loop functions (or ceases)* in those different states—we are actually talking about two little walnuts tucked deep in our brain: our thalamus. This symmetrical pair of nuggets that rests atop our cerebellum is connected to myriad parts of our brain—serving as a destination, way station & departure point for the multiple interweaving loops of data, experience & response that turn the gears of our consciousness.

Generally speaking, all chordates (the lineage that has led to spinal columns—from invertebrates like sea quirts and hagfish to vertebrates like sharks and everything after) employ 6 primary systems to manage nerve-based "data & response" behavior. An external data input system (external senses), an internal data collection system (internal gauges of temperature, etc.), external & internal data analysis systems (to determine responses), a motor control system (to enact responses), and a *switchboard* (to help manage responses). In vertebrates, the thalamus & basal ganglia essentially serve as the brain's two primary switchboards—receiving data from their network of connections and using neural circuitry & neurotransmitters to help inhibit & enact a vast variety of scripted responses in other parts of the brain, and ultimately in nerves & muscle fibers throughout the entire body.

The more evolved the brain, the more complex the role of these switchboards in handling all of the neural traffic that generates the seemingly uncountable aspects of behavior & cognition. This means that in the ultra-evolved brains of humans, the thalamus has a long, broad list of responsibilities. For example, all of our sensory systems (except the olfactory, which is handled by the basal ganglia) route their data through the thalamus in its journey through the loop.

The biggest switch controlled by our thalamus is one that relates directly to dreaming: the *conscious/unconscious switch*. Without a functioning thalamus, consciousness cannot be achieved (i.e., damage to the thalamus can result in coma).[5] Because it's believed that the thalamus switches our brain from waking mode to sleeping mode, it's likely that the thalamus is also the key that starts-up our dream-state consciousness. Research has, indeed, shown that during dreams parts of the thalamus are active, and its connections to portions of the cortex and the visual system are also engaged (& interacting in uniquely *associative* ways).[6] In this dream-state, it is therefore also likely that our thalamic "switchboard" is what helps to regulate that altered loop of consciousness—e.g., switching off (or not awakening) those validity mechanisms used in the building & analyzing of narratives.

Evolutionarily, it makes sense that between our brain's two primary switchboards—the thalamus & basal ganglia—the former came to control our modes of consciousness. Although both first appear in ancient vertebrates like sharks, and the two organs communicate reciprocally (inhibiting & enacting across their wires both competing & cooperating motor scripts originally generated via all their data input sources) the thalamus & basal ganglia, as noted, controlled distinctly different sensory switchboards: visual & olfactory, respectively.

As brains evolved, more advanced animals began to rely much more heavily (& complexly) on vision & the other senses managed by the thalamus. This ultimately lead those early cerebral cortexes (heavily-founded upon expanded, complex optical lobes) to become the engine of cognition in mammals. Thus positioning our visually-rooted switchboard to take over as the manager of the most important evolved feature of those magnificent vertebrate brains—the feature that developed a clever on/off & *on-but-only-dreaming* switch: consciousness.

Why The Weirdness Is Drawn in Crayon

As mentioned, science has suggested that our consciousness (via our thalamus) makes use of our visual systems while we're dreaming. And although we usually spend our time in these essays worshipping & praising the brain's ability to create beauty, there's something slightly unkind that needs to be noted about our consciousness's use of those visual systems, about its skill as a solo "set-painter." It's something that becomes fairly clear in our dreams: left alone, our consciousness is, frankly, *a hack*.

Our brain's real virtuoso creator of settings—its union-sanctioned cinematographer—is our *eyes*. But that part of our system is entirely dependent on actual incoming environmental data. Our memory-stored visual data simply isn't as rich & detailed as the real stuff consumed via those ocular organs. Thanks to our data limitations, in its memory-stored form, we only get the *general contours* of that original incoming visual data when using it for association, comparison & recall. (Another matter we'll explore in detail in our next essay.)

Therefore, when we're building the settings of our dream worlds, the *general contours* of our environment are all we get. You might think of this dream environment as a *low resolution* version of the real world. That incredible detail in a waking moment requires vast piles of visual data, resulting in a highly-detailed moment-to-moment rendering. With less rich visual data to build from, our dreams can only achieve a medium or low detail setting. Which means that dreams can still contain vivid, colorful, and powerful images—but ultimately, each individual dream scene or moment is limited in the quantity & quality of what it is depicting. And this leads to these scenes usually containing a handful of dominant & powerful elements, but a kind of *fuzziness* around the edges or in the more fine-grained detail of the situation.

What The Weirdness Tells Us

Of course, whenever we talk about dreams, the real question in the back of everyone's minds is the same, and it has nothing to do with image resolution: *What do my dreams mean*? The answer: that depends. Frankly, they don't usually *mean* much at all. In other words, they aren't *intended* to tell us things that we somehow can't express to ourselves consciously—our dreams don't have a *strategy* for what they're telling us other than essentially random association to current high priority data & the narratives that flow from it.

Nonetheless, we might still *learn* things and *infer* useful information from our dreams both consciously & subconsciously. The primary "truths" that our dreams might tell us are which recently experienced events or contemplated ideas have been perceived as fairly important during the course of the day, and thus have influenced our dream narratives, which are more likely to use that recent higher-priority data during the process. Or they might foster & reveal connections between ideas or events that are sneakily-related but were not previously linked by your day shift consciousness.

And through those new narratively-created applications & associations of stored data, our dreams are capable of opening possible pathways to otherwise unlikely-to-be-generated ideas or actions—which is really their main purpose.

Why, Exactly, Must We Watch All This Night Shift Weirdness?

So then, our last question about dreams echoes that central question we asked about our internal dialogue loop: why do we actually *experience* dreams? In other words, when our sleeping brain is running its little super-fluid data application & association night shift sub-program, why does it bother to tell our consciousness this silly story to go along with the sub-program? *It can't help itself*. It's the only way our consciousness works. Our brain can still run the narrative loop without external input, and that loop can skip a side-station like validity (which recent science is suggesting happens in a very specific region of the brain). But that narrative-building, thought-generating, inside-your-head experience is a *necessary part* of our consciousness' looping generative & associative process.

Once we've closed off our external input data pipes, that narratively-built internal dialogue & accompanying experience are the only possible sources of incoming data for our subconscious to process. As discussed throughout these essays—in our looping model, this sustained flow of incoming data is what *drives* those associative processes within our stored data. We are conscious in our dreams because we *need* to be in order to run the night shift's sub-program. As with every product of consciousness—our dreams *must* result from its narrative-loop internal-dialogue-based system, because it's the system that's required to access & apply *any* of our memory-stored experiential data. More specifically, our internal dialogue is the system that's required to access & apply that data *linguistically & syntactically*—which (as we'll discuss in our next essay) is how that experiential data is primarily *stored*, and thus the best way for us to broadly apply & associate that data.

Think of it this way: many pre-language mammals also probably experience dreams and use those dreams to generate random (& possibility-broadening) applications for & associations within their rudimentarily narrative, but purely sensory-based, experiential stored data. However, unlike humans' *wacky-yet-still-narratively-driven-&-complex* dream state, it seems likely that (by applying those rudimentary narrative structures to purely-sensory stored data) pre-language mammalian dreams sequences are less elaborately plotted & more narrowly focused in their scenes (yet still similarly weird & randomly-connected). *Chasing a cat around a circular, dirt-filled swimming pool, trying to eat kibble from a giant, wobbling bowl, pawing at a donut rolling along an endless park bench.* In other words, if we *didn't* power our own night shift via our consciousness-generating internal dialogue, and merely used purely-sensory data—our dreams would have about the same limits as *doggy dreams*.

In the end, every aspect of our consciousness is generated by these nested systems. Repress external sensory input & physical action responses, close the off-ramp to our validity station, spark up a recent (or high priority) piece of data in the cue, and *voila!* Those nested systems of consciousness now produce the night shift: the universe of our dreams. Ultimately, this *sub-waking* universe's purpose is, at its core, the same as the purpose of our internal dialogue loop in the waking universe: *to make us better, more adaptable problem-solvers.* To enhance the neural pathways that fuel the human brain's majestic ability to achieve unique & creative solutions to our most vexing challenges.

In the most romantic terms, our dreams—in the most romantically unbound & mercurial of ways—help show us what is *possible*. These doppelgängers of our mind's darkness are engaged in the tireless, shameless task of exploring the strangest & most unlikely paths through our consciousness, seeking in those folds of gray what previously-unexplored or least-traveled trails might lead us to the next beautiful idea that humankind has yet to have seen.

###

Endnotes:

p. 86:
1. Hobson, J. Allan. *Dreaming: A very short introduction*. Oxford University Press, 2005.

2. Hobson, J. A., and K. J. Friston. "Waking and dreaming consciousness: Neurobiological and functional considerations." *Progress in neurobiology* 98.1 (2012): 82-98.

p. 88:
3. Hoel, Erik. "The overfitted brain: Dreams evolved to assist generalization." *Patterns* 2.5 (2021): 100244

p. 91:
4. Hoel, Erik. "The overfitted brain: Dreams evolved to assist generalization." *Patterns* 2.5 (2021): 100244

p. 94:
5. Schiff, Nicholas D. "Recovery of consciousness after brain injury: a mesocircuit hypothesis." *Trends in neurosciences* 33.1 (2010): 1-9.

6. Chow, Ho Ming, et al. "Rhythmic alternating patterns of brain activity distinguish rapid eye movement sleep from other states of consciousness." *Proceedings of the National Academy of Sciences* 110.25 (2013): 10300-10305.

You Remember You

The Sum of Ourselves

Who are you? It is easy to identify *what* we are—this collection of muscle fibers, neural tissues, skeletal frameworks, this flesh and bone—but gleaning *who* we are is an elusive, mysterious, metamorphosing, and almost unfathomably-complex process of perpetual self-building that encompasses every moment of our conscious being.

And what is at the core of this self-building machinery? Memory. The remaining sum total of everything that you have ever seen or done, every experience your mind has consumed. And what are these memories really? Data. Ginormous, explosively and exponentially interconnected, magnificently vast piles of data.

I'm tempted to describe our data's vastness as *incomprehensible*, except that this is exactly what the human brain was built to do: comprehend that data. Not only comprehend it, but decide how and when to make use of it in our moment-to-moment decision-making process. And for consistency, we'll say yet again: that is the ultimate purpose of our consciousness—to make lots & lots of decisions, every second of every day.

In essence, much of our cerebral cortex functions as our brain's data storage system, our *hard drive*. And while humans have been working on computer hard drives for mere decades, the forces of the universe have spent several hundred million years perfecting the technology that is our lumpy, folded, gray matter. Which leads us to ask: what provocatively brilliant solutions has the universe stumbled upon during the evolution of the human brain's sophisticated data storage systems?

When we consider these storage systems of the mind, we are also necessarily considering the systems' handling of data retrieval, comparison, analysis and application (essentially, our *cognitive* processes). Thus, any full blueprint of this data-handling machinery must depict a complex, dynamic architecture capable of adapting to the myriad short- and long-term challenges the brain encounters.

Despite this architecture's complexity, by applying what we know about the brain and our own experiences, we can hypothesize a set of fundamental memory & cognitive systems & mechanics that can help to explain the inner workings of our brain's hard drive & our language-based cognition. Narrative Complexity seeks to do this by exploring how the brain stores & handles memories & thoughts in their most familiar and fundamental form: as narratives.

(There are, of course, other kinds of non-narrative "task-based" memories—those myriad, detailed & deeply-remembered motor scripts that we use to physically enact everything from walking & eating to hitting a baseball. But those kinds of learned physical scripts are stored in different ways & locations in our brain, and are handled by those more primitive & essentially *unconscious* mechanisms of our pre-human systems of mind—which we'll explore in our final essay.)

A Story From Your Life

The simplest way to view these narratively-based memory mechanics is in their natural habitat, to trace their workings within the environs of everyday experience. *You are running along a familiar trail in the woods, a route in which every dip & turn has already been memorized. Suddenly, you encounter a freshly-fallen tree crossing the path. It stops you in your tracks and requires careful negotiation.* This is important, relevant, novel & valid data (yes, there's that omni-present *Narrative Prioritizor Test* again). In other words, we should probably *remember* this.

Throughout the entire run, you've been taking in environmental data & matching it to previously recorded data about the path, using it to help guide your course and pace based on your resources and goals. But it's likely that on any particularly average day, your familiarity with the path combined with a preoccupation over other life-matters might lead your brain to neglect recording most of that non-novel environmental information while it focuses on processing internal dialogue narratives about those specific life-matters. In these cases, upon later recollection you will likely have a memory of what you thought about, but not the specific details of, say, the trail's dampness. (Unless that dampness, for example, made something along the path reflect in a unique or beautiful way, causing that momentary image to attach to any internal narrative in which your consciousness was engaged.)

This focus on life-matters dialogue, however, can be interrupted when you encounter something along your run like a suddenly-narratively-important fallen tree. Now the "story" of our trail run takes precedence over the domestic drama in our head. In essence, the vehicle carrying our life-matters has exited the internal dialogue highway, and the typically-low-priority vehicle transporting the story of our run has sped onto the main thoroughfare. And it is the stories that occupy this prestigious roadway of our consciousness that are candidates for actual recording in our brain's memory database.

For the same reasons that we can't focus our conscious awareness on *everything* in our purview, there is no way we could *remember* everything that happens around us. That's just way too much data. And it would be essentially useless, because if we're recording *everything* then we're likely not prioritizing *any* of it. Once we begin prioritizing, there's no need to record all of it—because we can discard what isn't important, which is a much more efficient way to handle data. As soon as we accept that our brains must prioritize information in order to make use of it, it seems we must accept the likelihood that it would make use of this prioritization in selecting what information it records. A similar "prioritization mechanic" also helps determine which of those temporarily-recorded memories survives long-term.

And we need look no further than our life experience to find abundant evidence of this. We are more prone to remember specifically important, relevant, novel & perceivably valid moments or narratives over those that we judge to be insignificant, irrelevant, redundant & apparently unreliable. *Of all the mornings you drove to work that month, the only one you remember was the morning when you ran the red light and almost got hit by another car.* This mechanic is so obvious, examples almost seem superfluous.

The method that our brain uses to encode data with and calculate this prioritization is the system explored in our second essay: emotions. When a thought & its correlating experience enter our subconscious for recording, association, and subsequent thought generation, they're accompanied by the emotions generated when that thought parcel was first built. Those emotions were initially used to help guide the resultant actions and/or behavior. But once that's happened, these emotions serve their other purpose: to help encode & prioritize the newly-stored data & strengthen any associations it creates with other memory-stored data.

Of course, although these mechanisms work as a loop, the process can happen so quickly (in less than a second) and repeat with such extraordinary rapidity that it feels instantaneous to us. We can construct a thought parcel, then feel, perceive, store, associate, compare & evaluate its data seemingly all at once—running the loop of our consciousness in a snippet of time more mere than a moment. The deeper we dive into the mechanics of consciousness, the clearer the connections become between these nested systems of the mind—further revealing the elegant way in which all of its mechanisms are enacted and interwoven through our ever-efficient, perpetually-circumnavigating internal dialogue loop.

Memory Construction: Sentence-by-Sentence

So then, what really happens to this story after it makes its pass along the roadway of our consciousness and enters our data storage system? Essentially, that narrative information—the linguistic elements & syntax of that experience's correlating internal (or spoken) dialogue, plus the temporally-simultaneous, attention-defined associated environmental, physical & emotional data (sights, sounds, smells, tastes, sensations & feelings)—is recorded in our brain's neural network. More specifically, it's electrically and/or chemically imprinted onto those amazingly complex, interconnected, modularly-capable, differentially-associated, programmable & *re*-programmable neurons that compose the parts of our cerebral cortex that store data. In typical brains, these memory/data recording, recalling & associating mechanisms (our *parallel processing*) appear to primarily occur in our right hemisphere.

This is one of the areas where our model diverges from most current theories on memory management—because most suggest that memories are first stored & processed by the hippocampus before being transferred to long-term memory networks in the cortex. As we'll explain later—in our discussions of "short-term" & "working" memory—any kind of "intermediate" storage system would be an inefficient and ultimately *unnecessary* cognitive mechanism when viewed within the context of our model. As we'll also explain later, we hypothesize that the hippocampus clearly *does* have a specific & significant role in the formation & storage of new memories (as suggested by the unique memory deficits demonstrated by patients with damage to their hippocampus) but we propose that the primary memory data is initially encoded into the *cortex*—with the *aid* of the hippocampus.

And in 2017, researchers revealed the first evidence that mammalian brains do, indeed, encode new memories in the cortex at the same time that the hippocampus is handling these new memories [1] (thus contradicting previous models and, *theoretically*, supporting ours). Although our model *does* contradict the study's conclusions about how & when those cortex-based memories are used, I believe that's mostly because the study was done on *mice*—mammals who were among the *very first* creatures to employ cortex-based memories, and thus likely only present us with a highly *vestigial* view of humans' far-more-advanced modular memory systems.

The *experience* that is the source of this non-representational (aka, component-based & not truly *movie-like*) memory "recording" is what we perceive to be our true *in-the-moment* consciousness: the briefly-sustained, temporally-united experience of internal dialogue parcels combined with sensorially-perceived environmental & internal physical data that produces the *essence* of each moment's conscious experience. The locus of this process appears to be in the prefrontal cortex, but is, in total, a dynamically-constructed & ongoing effect of the simultaneous activation & integration of multiple distinct networks; this view mirrors the neural model of consciousness presented by Gerald Edelman's *Dynamic Core Hypothesis*. [2]

Since we've mentioned Edelman, I'll pause here to note that Edelman's & Tononi's *Universe of Consciousness: How Imagination Emerges from Matter* (Basic Books, 2000), Terrence Deacon's *Incomplete Nature: How Mind Emerged from Matter* (Norton, 2011) and Peter Ulric Tse's *The Neural Basis of Free Will: Criterial Causation* (MIT Press, 2013) together help provide a foundation for the neuroscience-based assumptions that underlie the mechanisms & systems I propose in this essay [3, 4]. Because I developed my initial hypothesis before actually *reading* those books (the latter two hadn't even been published yet) most of Narrative Complexity*'s* systems were not originally built upon the specific neural theories & mechanics that those books present.

My original proposition was that such neural mechanisms *must* be present (based on applying our theory to current knowledge of brain anatomy, behavior & evolution) in order for our model to function as theorized. And in the brief time since developing my initial hypothesis, Deacon & Tse have added to the neural evidence presented by Edelman—demonstrating that many of our theory's required mechanisms likely *do* exist within the human brain. In addition, the ideas in Douglas Hofstadter's *I Am a Strange Loop* (Basic Books, 2007) strengthen my conviction that a self-generating & self-sustaining looping stream of data (*language*) flows through & unites all those mechanisms in generating our consciousness.[5]

Nonetheless, our systems here will not be presented within the specific contexts of Hofstadter's, Edelman's, Deacon's or Tse's work (detailed explanations of which would make this essay lean too-heavily away from our *behaviorally-based* depiction of how these systems define human experience). Our goal here is to explain how the more general neural *capacities* & *abilities* that we hypothesize allow for the mechanisms that our model proposes (and anyone more interested in those *detailed explanations* that support those *capacities* & *abilities* is highly encouraged to read those four amazing & absolutely vital texts).

Swinging back around & returning to those actual *memory/data recording systems*... Current evidence also suggests that, in addition to our cortex-based memories, our amygdala is involved in storing & responding to specific kinds emotional memories [6]—primarily intense pain- & fear-based ones.[7] The amygdala's involvement in managing these kinds memories likely began in reptiles. Although our modern amygdala has its roots in the original vertebrate version (belonging to sharks & jawed fish) recent research has

shown that amniotes—reptiles & mammals—provide the first evidence of the amygdala developing sensory-associative regions. (Prior to reptiles the amygdala is limited to managing responses to internal data.)
In the view of Narrative Complexity, this amygdala-based memory is our most-ancient memory system, and might even be seen as an almost vestigial mechanism when compared to our other—primarily right-hemisphere-based—modern memory systems. This amygdala-based memory mechanism also appears to work essentially subconsciously & more reflexively than our primary memory system.

We can see this kind of subconscious operation in a famous century-old experiment by Édouard Claparède involving a woman who no longer had the ability to form new memories. Even though she could not remember meeting anyone new—no matter how often they would meet—in the experiment, she nonetheless recoiled from Claparède's attempted handshake the day after an encounter in which his handshake had included a painful prick. Not only did she not consciously recall the painful first encounter, she couldn't explain why she recoiled—to her it was simply a reflex.

As noted, research indicates that these specific kinds of reflexive pain- & fear-based memory responses are managed by the amygdala. Since our century-ago amnesiac's problem was in the formation of those new right-brain "conscious" memories, she still had the *reflex*, but not the *recollection*. It seems that the amygdala's was the first real memory system to evolve along the chordate pathway to the human brain. It was a kind of memory that allowed creatures to make use of unique remembered (& likely mostly pain-based) experiences long before the primary mechanisms of modern human consciousness (such as our cortex-based, right-brain memories) had begun to make their magic.

Returning to our primary, right-brain memory systems, as soon as a new parcel of internal dialogue (a sentence or phrase of a thought or an idea) is laid into the neural network—creating the foundation for a potential long-term memory—its component parts (specific words, images, emotions, etc.) build associations (synaptic pathways) to related, previously-recorded data. This data-pinging neurally connects the new potential memory to relevant (and ideally high-priority) stored data for current & future association and comparison. This data-pinging process also helps to produce the currently-percolating next thought, which will seek to use the highest-priority (most important, relevant, reliable) and/or most uniquely-applicable just-pinged data in constructing its next link in the ongoing narrative chain.

Link-by-link, our elements of daily experience—almost always set within or built around these narrative structures—sear themselves into our data-recording neurons, connecting these potential memories' modular elements in both a linear, syntactic, temporal fashion, and in an all-manner of all-angles vertical/diagonal, associative fashion. (Whether or not this potential memory becomes an *actual* memory partly depends upon the degree of *searing* when that narrative is laid into our neural network—something we'll explore in detail later.)

Who are you? In many very concrete ways you are simply & complicatedly a result of this process, a set of dynamic responses determined primarily by a lifelong chain of memories—a chain laid down one link, one moment at a time.

Our Outer Limits: Data Resolution

Although we've taken it for granted thus far, if we truly want to comprehend what's going on behind the veil of our consciousness—producing the many complicated facets of memory & cognition—we must genuinely consider the extraordinary *computational depth* of our mind's memory mechanisms.

The human brain is a massively powerful pattern association & comparison machine—subconsciously parsing a wide array of large & complex data patterns into their diversely modular components, then associating & comparing those components to related pattern data. The multitudinous elements and aspects of each self-contained memory-defining narrative pattern (and the attached environmental & associative data) are quickly examined & compared by our subconscious with a depth, detail & breadth that we are only minimally aware of consciously.

Generally speaking, we're only consciously aware of the emergent result & some feeling of the nuance behind these powerful calculations. But those flashes of neural activity—instantaneously circulating through the maze of memories & patterns, matching their ones and zeros against synaptically-connected stored data, and helping bring forth to the stage of our consciousness the thoughts we *perceive*—that perpetual lightning-storm in our brains goes primarily unnoticed by our conscious mind. Forever rapt by our moment-to-moment stories—which are the only things our minds were built to consciously comprehend—the whirring and sparking of the engine that builds those stories remains hidden in our neural silence.

Now that we have considered this, we must acknowledge an opposing truth: despite its deep & highly-complex ability to record & compare these memory-stored data patterns, the human brain is not *infinitely* powerful. This means, among other things, at some point its memory systems reach the limits of something we'll call *data resolution*.

One of Narrative Complexity's central hypotheses is that these neural networks that compose our memory databanks function primarily modularly. This goes back to our first essay's discussion of early animal brains representing ideas with one "word" (a singular neural component) that only & specifically means "I saw a red snake by the river this morning." In contrast, human brains employ a collection of individual words (modular neural components) that are combined to represent the same idea in several distinct, but malleable & independently-associative parts.

Basically, using a larger number of modular components to construct a full idea likely allows each component (and the full idea) to have greater *data resolution*—more capacity for informational detail—than when constructing the same idea using fewer (or a single) neural component(s). In essence, the latter method stuffs more items or pieces of data into its neural component(s), thus limiting the informational detail of this data.

Narrative Complexity further hypothesizes that the "skeleton" or framework around which all memories are built is language-based. In our theory, memories are narrative pattern-structures built from modular word-based elements—elements whose meaning & functionality (aka, their associative & generative cognitive capacities) are primarily defined by the words' broader symbolic content combined with their more specific linguistically-defined semantic content and syntactically-defined functional roles. Basically, these narratives that ultimately compose our recorded memories are built upon & around the words that we say to ourselves as the experiences occur (& the words we use when remembering or retelling the stories).

The modularization of these narrative parcels (a story, sentence, event sequence, etc.) that compose a memory is critical to data resolution. That's because our *memory modules* (those component parts of a narrative parcel) ultimately have a defined data capacity and a defined capacity for external associative connections. This is obvious because otherwise, the power of our memory & associative capabilities would essentially be infinite. Thus, as described earlier, the more individual pieces of data that are recorded onto one narratively-determined memory module, the lower the resolution of each piece of data's informational detail.

These *memory modules* are likely a collection of neurons arranged in (or a single neuron containing) a standard, defined structure that represents an individual memory module. We might imagine (to view the structure overly-rudimentarily, but in way that makes this concept easier to visualize) that our "short-term" memory's typical 5-7 item limit (which we'll discuss in a moment) is actually a reflection of the number of neurons that compose a standard memory-module structure.

It might seem *fundamentally odd* that our brain would pre-select a specific number of neurons to compose such a structure. This is because it makes our brain feel like a system designed by someone who contemplated choices: *"Hmm… how about we try using 6 neurons for a module. Create a model using those variables…"* But, in essence, this is what the process of evolution is doing. And at different points in evolution different "test-models" become stable for certain periods in a species. Very recent research, in fact, shows that our visual-spatial systems have made these kinds of oddly-specific-seeming *choices* in how it manages data. For example, it turns out that our brain spatially "grids" our world around us using *triangles*.[7] Why not *squares*? Or those awesomely-inter-connective *hexagons* that comprise all of the hippest board games today? Well, because that's what evolution has settled on in humans *for this cosmic moment*.

Similarly, somewhere amongst those evolutionary algorithms, our brain has arrived at some standard, pre-defined neural structure (limiting capacity & associations) for individual memory modules—which map to our use of syntax in narrative/prediction-construction. Basically, according to our theory, each individual module would correlate to (& contain) a single word or a small, limited group of words that make(s) up a specific syntactic element (e.g., subject, object or predicate, etc.).

One place where we can see this *memory-module-capacity-&-associations* mechanic at work is in the use of memory devices that aid in recalling data like lists. In my freshman year psychology lab, we did a simple short-term memory experiment. Or rather, *what is often mistakenly described as short-term memory*. As we'll make clear in this essay, in the view of Narrative Complexity there is no intermediate, quickly-disappearing "cache" of recently-consumed data—aka short-term or *working* memory. In our theory, the effects of short-term memory are explained entirely within one all-encompassing data-storage mechanism.

Which brings us back to that psych lab. Our instructor listed single-digit numbers out loud, and we were told to remember as many as we could in order. The experiment was meant to demonstrate our short-term memory's typically-limited capacity to contain a list of individual items. Much currently accepted theory suggests that a typical short-term memory has (as mentioned) a limit of 5–7 items—a theory that was reinforced that morning by almost all of the 15 or so students, most of whom recalled 5–7 items.

But one other student and I were able to remember a lot more, each of us recalling about 14 numbers. Both of us used simple memory devices to aid our recall. I—a devoted Chicago Bears fan—had gotten into the habit of pairing all numbers and remembering them according to a corresponding Bears player. Thus, I wasn't really remembering 14 individual items. I was still remembering essentially 7 individual items, but each was capable of associating itself with previously-recorded memory data that already contained a 2-digit numerical component.

Basically, I was maximizing the use of my memory module's limited data capacity by employing its items to access data stored outside of itself via associative connections. In the other students, the memory module containing the list used each of its (imagined) 5-7 neurons to record a number, and I used each of those neurons to record a name linked to already-stored data. This kind of memory device is often referred to as "chunking" data—which is a misleading label, because (as we'll show here) we're *always* "chunking" data in some fashion or another. (The other student, by the way, used a visual/tactile device imagining dialing numbers on a phone, which is another version of what I prefer to call *data maximization*.)

Now consider this memory challenge: recall the names of 20 people immediately after learning them. Again, most people would struggle to remember more than the list of 5-7 that seems to be the "item limit" of our brain's individual memory modules. And without the numerical component, my simple Bears device would not allow me to maximize my limited data-capacity here (*no "chunking"*). But a memory expert might create and tell themselves an internal story when learning the names. Immediately afterward, once they recall the first name & set the story in motion, the rest come tumbling forth.

In most people's minds, their basic narrative here is something like: "I am remembering a list…" or maybe a simple self-directive "Remember the names…" *Boom. You just lost the memory game.* The recall-resuscitating syntactic logic/structure of the narrative hits a dead end at the modular syntactic element: "a list" or "the names." That *list* or those *names* (whose occupation of a single module has been determined by how that linguistic element is being used in the narrative syntax) has now been deemed the dumping ground for all subsequent data that composes the list. In others words, the data capacity for the entire list has been limited to this single module, dooming the list to run out of space after it reaches the module's 5–7 item limit. Additionally, the data resolution of its items—which is essentially a result of the module's data & associative capacities being divided among those items—has also been reduced by stuffing the whole list into one memory module.

But the memory wizard has put the power of story to work. Not only have they turned the list of names into a modular narrative, but by devoting individual memory modules to each item (a result of the narrative syntax) they increase the capacity for each individual item's data resolution. Thus, in addition to being able to remember more names, they're also possibly able to remember a few specific details about each person in the list. (Another memory-enhancing technique—creating a visually-oriented "memory palace"—is essentially a spatially-based narrative, if you think of the individual rooms like "scenes" & the sequence/layout of the rooms like a story structure.)

When we try to remember something like a list, the way in which we *syntactically* (according to our narrative) construct or perceive that list helps determine how those pieces of data are modularized, and therefore how much data capacity & associative capacity is available to each syntactic element (& its components) for recording.

Instead of stuffing lots of data into individual modules (like those early brains) narrative helps us to make use of associative connections *between* memory-stored data modules. Basically, our brains are designed to follow everything that pops into our head with the compulsion to complete imperatives like *and then…*, *then why…* or *because…* in order to help build a cohesive story. (If you don't believe this, try talking to a 3-year-old, whose *imperatives* have just started to emerge.) And narratively logical construction (valid pattern-building) helps imprint a sequences in our memory.

Our brains work this way because these kinds of narratives are deemed *highly valid*—especially reliable as predictive patterns. In other words, this data is arranged in a way that adheres to learned rules of causality (rules that govern both linguistic & narrative syntax/prediction) and follows the contours of known, related & reliable patterns. These are the predictions that our consciousness was built to make in response to all of this incoming data. Inserting 20 names into a coherent & engaging narrative allows the brain to escape the limitations of an individual data module's defined storage capacity and make use of its more powerful & efficient narrative & associative abilities—to take full advantage of the human brain's magnificently-evolved neural modularity.

The Volume of the Voices

For me, one of the hardest truths to perceive when deciphering Narrative Complexity's explanation of consciousness & its attendant mechanisms: we're actually *always* talking to ourselves or someone else. Or we're fully engaged in *external* dialogue—someone else speaking, a movie, a book, *this essay*, etc.—which can essentially substitute for our own internal dialogue, an experience enhanced through empathy (discussed in my Story Theory essay).

It's true—we narrate *everything* in our minds. Even when you try to "catch yourself" not *thinking* internally, just listen—there you are talking to yourself about how you aren't thinking about anything. It's essentially impossible to "hear" yourself *not* thinking for more than a few moments (if that). Just try it. You can't. (Then *stop* trying, because it'll start to drive you nuts—trust me.)

I know what your internal dialogue is thinking right now: *what about meditation?* I've never been a Buddhist monk, so I can't speak to what internal dialogue manipulations they have either achieved or fooled themselves into believing they've achieved. But in your average, everyday, enlightenment-seeking, yoga-mat-toting suburbanite or city-dweller (which I *have* been on occasion) I can tell you what is likely common among all of *us*. Even if you are trying to think of *absolutely nothing*, your brain cannot comprehend the true *absence of everything*—at the very least your nothing is pure whiteness or grayness or blackness. And, like it or not, the "sight" of that color in your mind generates the related word—*your brain can't help itself*. Sure, "white" is a pretty vapid thought, but it's still *internal dialogue.*

(This is not to entirely *dismiss* the actual physical & mental benefits of meditation, which—without going into the matter too deeply—can ultimately result from quieting that cacophony of complex & often *stress-producing* internal dialogue by replacing it with a much more vapidly-serene, simple & purely-sensorially "experiential" internal expression of *white.*)

Despite its perpetual nature, some of this self-narration is built from such brief rote scripts and mundane elements (*It's darker. That's crooked. Where did I put that?*) that we likely barely *notice* the words spoken in our heads—either because the thought was experienced almost instantaneously and/or it was of such low priority that it was essentially a whisper along the roadway of our consciousness. And I mean *whisper* in an almost literal way—this is because internal dialogue appears to make use of our auditory cortex as it emerges in our consciousness (and dysfunction within this internal data exchange can result in auditory hallucinations [9]—essentially, falsely ascribing products of our internal dialogue to outside sources).

Thus, Narrative Complexity hypothesizes that these lowest priority thoughts are processed by our auditory cortex like an actual whisper. Obviously, there is no literal "volume" to this kind of internal dialogue experience, so what does this really mean neurally? Consider that in terms of processing external sound data, the auditory cortex produces different results within our consciousness mechanisms according to volume. Loud noises are more likely to garner our attention enough to spark & perpetuate their conscious contemplation than very quiet, hard-to-notice noises. (Which is likely why music provides a more emotionally intense & immersive experience when played *very loudly*—an experience that, admittedly, I have great personal affection for.) This kind of *differentiated attentional response* is essentially mimicked by low priority (internally "whispered") & high priority (internally "shouted") dialogue—which makes that louder/higher priority internal dialogue more likely to garner our attention enough to spark & perpetuate its conscious contemplation than quieter/lower priority internal dialogue.

In fact, the only reason that quiet, mundane, low priority thought even *reached* the conscious roadway is because our brain appears to grade narrative priority *on a curve*. This seems to be one of the effects of perpetually needing to narrate our lives. Something *always* has to be running along that roadway, so if *every* current potential narrative vehicle carries little weight, the heaviest of the little gets its chance to ride the open road. When it gets there, the scene is like 3 AM on a lost highway in the plains—so quiet that its wisp of a narrative almost gets noticed by you, but maybe not. Or at least not until you find your nearly-silent self in the open roadway and are prompted to ask: *what was I just thinking?* You might be prone to answer *nothing*. But that wouldn't quite be accurate—in truth, you were thinking *something*, but the thought was barely worth hearing. Or remembering.

There are actually some common techniques in which we naturally adjust this never-ending internal dialogue to help with memory-management. For example, when someone tells us to "hold that thought" we might instinctively try to *slow down the pace* of our internal dialogue or *repeat to ourselves* the thought we've been instructed to "hold." Both techniques are different ways to prevent ourselves from laying down new narrative parcels into our memory; this ensures that the "held" thought is the most-recent (thus, an easily-accessible) piece of data in our storage. Additionally, the latter technique (repeating) also helps to strengthen the data's imprint. (These memory imprinting mechanics will be discussed in detail in the next section.)

The mechanisms within our loop that permit some narratives threads to emerge in our conscious awareness while other (essentially simultaneous, but currently less-prestigious) potential narratives remain confined to our unaware subconscious is a concept that Narrative Complexity refers to as the "Diffuse Box of Consciousness." We'll explore in detail the workings of these mechanisms in our next essay (such as how those potential narratives subtly affect our behavior & decisions despite our conscious unawareness of them). In terms of our memory, the most significant result of this *Diffuse Box*: only the narrative threads that reach our actual consciousness can be seared as a memories.

And the threads that weave their way onto our conscious roadway essentially travel in one of two kinds of vehicles: "spoken" and "experienced" internal dialogue. As we just explained, those low-priority or instantly-fleeting conscious thoughts are not always "heard" by us word-for-word in the way that higher-priority or more-deliberate, focused internal dialogue is "spoken" inside our heads. Nonetheless, our speedy or prestige-challenged "experienced" internal dialogue is not totally invisible within our loop. Because the dialogue made it onto our conscious roadway, that word-based thought parcel still takes the narrative path into our subconscious processing—where it is weakly seared and meekly aids in building the subsequent thought. And its (likely short) existence as a potential memory is why this mundanity is still briefly available for immediate recollection in response to the sudden "*What was I just thinking?"* prompt. Of course, since this dull data is essentially the weakest form of a potential memory, it doesn't often linger for enough time *to be* remembered, thus leading it to be forever lost.

How to Make a Memory

Which brings us back to that matter of *potential* memories becoming *actual* memories. Even though we are always talking to ourselves, and those narratives become the foundation (or at least the starting point for the foundation) of long-term memories, we obviously do not remember every single sentence of internal dialogue.

In fact, you probably can't even recall most of what you said *out loud* during breakfast this morning. And yet, if one of those sentences was a response to your partner announcing "I'm pregnant," then it's likely those sentences and some of the moment's surrounding details would be well-remembered. *So how does that happen?* How does one set of sentences become a long-term memory while other dialogue is entirely lost? In the view of Narrative Complexity, the three key players in this mechanism are emotion, repeated recall & recentness.

As we've described, when any narrative parcel flows from our internal dialogue into our subconscious processing, that language-based pattern is seared into our data-storing neurons, becoming the foundation for a potential long-term memory. According to our theory, the degree of searing is mostly determined by the level of the specific emotions that came attached with the narrative parcel. Powerfully emotional narratives (like someone saying *You're pregnant!?*) are initially seared with commensurate power, while mundane narratives (like someone saying *I couldn't find my razor...)* are given a weak initial imprint. This immediately makes those higher-priority, more-powerful narratives stronger long-term memory candidates. In fact, they are already likely seared in a "semi-permanent" fashion—but whether or not that memory grows stronger (becomes "permanent" or much more likely to be recalled in the future) is primarily dependent on another of those key players: repeated recall.

This repeated recall is also essentially the only way that a mundane memory can become a long-term memory. This is because there is also that *third* key factor affecting our data's imprint strength: its *recentness*. Narrative Complexity hypothesizes that the most-weakly initially-imprinted potential memories have, essentially, a very short "half-life." In these cases, the only thing the memory has going for it is its *recentness*—once the memory is no longer recent, it's likely no longer there.

When a memory has enough emotional juice, it seems to immediately exceed that minimum level of searing below which a super-quick version of the memory-degradation countdown begins (thus placing the initially stronger memory into that "semi-permanent" state). But when a memory has none of that emotional juice and is seared only by the minimum level of neural energy provided by passing through our dialogue loop, that memory is immediately counting down to soon-to-be oblivion. If this data is not quickly accessed again, the strength of its imprint soon fades from the neurons. (Likely returning them to blankness—making them available for future re-use until something is more-strongly seared there or possibly, if they've been "re-used" multiple times and somehow degraded, allowing them to be discarded eventually & replaced with new neurons.)

This aforementioned "emotional juice" & "neural energy" might be described more accurately as electrical and/or chemical energy that accompanies a thought parcel (or is simultaneously present within those cognitive networks) as the parcel passes through the loop, and ultimately represents that parcel's initial *memory potential* when seared into the data-storage neurons. This searing only *allows* this data to be remembered (either temporarily or long-term) and thus merely makes it *available* for future recall. Whether or not this data is *likely* to be recalled (essentially, to become a stronger or permanent memory) is heavily dependent on the mechanisms of repeated recall.

Once a potential memory or semi-permanent memory is laid into our neurons, this likelihood of future recall is primarily determined according to the paths by which and how often that data can be reached via future pinging. Part of this "how often" is determined by searing factors like the initial emotional priority given to high-impact memories, whose powerful prestige can automatically make them generally more likely to be more frequently pulled from our data pile when related data enters the system. (Post Traumatic Stress Disorder is essentially this mechanic forced into overdrive by extreme memory data—a problem that's also likely worsened by those additional, powerful pain/fear-related amygdala-based memories of the trauma, which trigger essentially subconscious responses.)

Another part of this future-recall likelihood is determined by the associative neural pathways that might lead to that data. When a potential memory is first seared and creates synaptic pathways to other data, the number of associations and the kind of data with which it associates both affect future recall. If a potential or semi-permanent memory creates a lot of initial associations to other data & if that data is high-priority, with lots of its own pathways to other high-priority data—that's a best-case scenario. This memory's links to lots of information that's likely to be pinged makes the memory itself more likely to pop-up in future thought-branches.

In addition, when one of those pathways to the memory is used, and thus "enhanced" by the traveling neural energy (basically, when the memory is pinged as related or relevant to incoming data) that pathway seems to become more *fluid*—more likely to be traveled again when related data returns. Imagine these flashes of neural activity running our data-storage maze via associative connections between memory modules. More-fluid paths act like broader, more-easily traveled neural roadways. Less-frequently accessed memories seem to have weaker or less-fluid synaptic pathways connecting them to other data. Like narrow trails leading away from well-traveled thoroughfares, these paths are more apt to be passed by this neural lightning, which is likely seeking the path of least resistance (greater *fluidity*).

And every time one of these memory-stored neural-networks is accessed, the memory data itself becomes a little stronger—benefitting from the newly-generated memory potential that has now reached this data via our narrative loop and those subconscious associative processes. This is why—as noted—if one of those mundane, ticking-down-to-oblivion recent memories is going to survive its half-life, it quickly needs to be accessed again: to strengthen or create new associative pathways and leech more memory potential from our dialogue loop to help strengthen the data imprint on our neurons.

Even long-term memories, of course, have a tendency to fade or degrade as they age, but it appears that those neurons' data imprint & association strength is on a much different chemical clock than mere potential memories. This would make perfect sense in terms of how the brain mostly needs to use this long-term data: for *future* analysis and decisions.

And it's logical that the *less* a long-term memory is accessed, the more likely it is to fade away. Recalling or pinging a memory is innate proof that it remains potentially relevant and useful—and every time a memory is pinged, the accompanying energy adds a little more time to its clock. If a long-term memory is *never* pinged, that's innate proof of its uselessness, and thus the clock continues its countdown unextended, slowly making its way into the brain's junkyard of the almost-invisible.

Do these long-term memories ever *truly* fade forever—their data imprints disappearing from those neurons like potential memories that never make it? That is a *very* difficult question to hypothesize about without being able to autopsy individual neurons. This is because there would essentially be no perceivable difference in the resultant effects from either the data disappearing entirely or the data weakly remaining, but the incoming paths becoming so impassible or buried far, far away from any likely-to-be-traveled neural thoroughfare that they are simply never accessed again. My *guess*? We've still got a lot stuff hidden in the attic, but we just can't seem to find it anymore in our hoarder-ish-ly overstuffed cranium.

Now You Have It, Now You Don't

Although most memory-recall events *usually* help us to more-easily access that memory again, under certain circumstances—in a quirk of our memory mechanisms—briefly recalling an old, weak piece of stored data can sometimes make it *harder to* recall the data again. These instances are reflected in those "tip of your tongue" experiences, when you are *sure* you know something (or just briefly, imperceivably thought of it) but cannot quite recall it. In these cases, the old memory was likely accessed through a "*uniquely* unique" associative pathway. In other words, *you only thought of it because… yada, yada, yada*. Usually *yada, yada, yada* is some weird of set of nearly-random-but-intersecting associations triggered by something unexpectedly.

In the midst of writing the e-mail, a scent wafts in the window: the aroma from a neighbor's dryer vent, an ancient olfactory experience from your childhood. At the same moment, you glance at a single yellow candle glowing in the dusk. This combination fleetingly recalls a specific, powerfully-emotional, but long-forgotten moment from youth that involved both details. For a second you feel exactly like you're there again, but then it's gone. Hard as you try, you cannot bring back the moment or feeling again, and cannot seem to say exactly what you were remembering.

What happened? We just remembered the data—so shouldn't that now make it easier to recall? Unfortunately, *not always*. That data wasn't recalled because it was connected to well-traveled pathways, it was recalled because of a "*uniquely* unique" association. When that old scent was pinged (and scent holds a powerful, primal imprinting capacity—which has helped keep this old memory alive & available) there weren't many pathways connected to the scent, and the fact that we were also processing another unique image that just happened to connect to that same specific childhood moment helped to create the perfect conditions for pinging that faint, hidden memory. (You entered *exactly* the right data into the search engine.)

The problem here is partly the memory's faintness. When this old data bubbles up into our subconscious, its weak remaining memory potential & lack of informational detail (due to that slow ongoing degradation & rare recall of this specific data) might hinder its ability to reach (or be properly depicted within) our actual conscious roadway. But even though the actual details of this old memory are trapped in our subconscious, its unique data-match still likely registers slightly in that specific part of our brain that gauges validity. (And its attached emotions are quickly felt.) This is, after all, a momentary successful pattern comparison—the incoming data pinged something that resulted in a successful cross-match.

Because our brain has matched the unique conditions of this moment with the conditions of a specific memory, it wants us to take note (feel) and compare the data to see if there's something we can cross-apply to this moment. In this case, there's nothing in that old data that applies to anything currently on the table. Nonetheless, even though this faintness & ultimate irrelevance means we aren't consciously made aware of the data's detail, we briefly *feel* the sensation of a successful match.

When this kind of *just-thought-of-it…wait-a-second* event occurs during something like a *Trivial Pursuit* game and the faint memory actually *does* pertain to a matter currently on the table—and yet we *still* can't bring it to mind again—the other contributing culprit here is likely the *fear & anxiety* produced by your powerful desire to look smart & take another turn. These kinds of "survival" emotions produced when the likelihood of success is low & stakes are high naturally inhibits the fluid exploration of neural data required to re-ping that faint & literally *trivial* data.

Once the scent is gone (or the initial flash of that trivial answer has sped away) we no longer have easy access to that unique neural pathway connecting the old memory data. And something else kind of annoying has probably happened. We've been saying to ourselves things like: "*What was that? The candle reminded me of something, and that smell. What was that?*" In other words, we've been laying down recent and possibly now-urgent-feeling potential memories that are associated with the same data-pinging elements that might lead us back to that faint, hidden memory.

In essence, we've created a closed loop in which trying to remember the lost moment is most likely to lead us right back to that now-more-prestigious, just-laid-down memory of thinking about remembering it. We've trapped ourselves away from that old data, which is too faint to butt-in on the maddening (and now repeatedly reinforcing) loop of "*What did that candle make me think of?*"

This is likely why the best strategy for re-pinging that *just-slipped-away* old data is to try "retracing your steps" back into the memory—to try recreating the specific mental conditions that initially led you along that unlikely backroad to the ancient, nearly-hidden piece of data. *We were talking about... then you said... and I said... and then the wind blew over those flowers... that's it—I've got it!*

Therefore, as we've described, the less-frequently we access old data, the harder it is to find (or stumble across) in our vast data pile. And when it does finally pop up, that old data doesn't often stay long enough to make much use of its brief cameo appearance (or we're too suddenly-nervous over a *likely & trivial -yet-ego-bruising* failure to fluidly re-retrieve that thing you just knew 2 seconds ago).

Another very familiar, common & weird *now-you-have-it-now-you-don't-ish* neural event: *deja vu*. In the view of Narrative Complexity, *deja vu* is an easy-to-explain yet hard-to-pin-down event. It's easy because there seem to be a range of ways for this experience to occur. One way: a "hiccup" in our data chain in which data traces faintly reactivate neural networks after departing, causing a "ghost" of the data experience to trail behind it, resulting in an essentially simultaneous experience that seems to "remember itself."

Another way: the "pattern match" emotional responses (like *affirmation*) that indicate direct correlations between incoming & stored/predicted data momentarily "over-express" themselves (either through a calculation error or a brief purely-neurotransmitter imbalance) resulting in a feeling of "over-familiarity" with stimuli or events.

As we're fond of saying, *in other words*, in a system like human consciousness—in which "reality" & our familiarity with a specific experience within it both essentially result from the *re-representation* of incoming data constructed in a link-by-link fashion—the question isn't *why do we experience deja vu?* The question is really *why aren't we experiencing deja vu basically all the time?* (The apparent & somewhat *unrevealing* answer: despite the seemingly strong likelihood that a dynamic, highly-complex system like human consciousness would frequently fall "out of sync"—amazingly, the brain does an admirably consistent job of mostly maintaining a fluid, *hitch-less* conscious experience. Or at least it consistently *tricks us into believing* the experience is fluid.)

The Illusion of a Short-Term Memory Cache

Confession: while you weren't looking earlier, we swapped out the concept of a *short-term memory cache* with our own *minimum memory-potential half-life* concept. Within the systems of Narrative Complexity, this half-life mechanism can explain most of the effects associated with a short-term memory cache. The other primary mechanisms that help explain these effects are the previously-discussed memory module capacity (which is the real cause of the "item limit" associated with a short-term cache) and the soon-to-be-discussed *narrative-building mechanisms* (which generate most of the effects associated with short-term memory's *handling* of the data in its cache, or *working* memory).

And since I've never actually done experiments on a live human brain to measure anything like the half-life of mundane data imprints on our neurons, it seems fair that I explain some of the reasons why I believe Narrative Complexity's system is more plausible and likely than a short-term memory-cache model.

We can get right to the heart of the matter by re-examining our memory-wizard's recall of the 20 names. In a system that relies on a separate short-term memory cache, what are they doing that allows them to escape the cache's defined item limit? Has the memory-wizard's use of narrative somehow expanded the actual data capacity of their short-term cache? Not likely.

If such a cache exists, its data contents must be limited in one of two ways. One, it is only limited temporally—meaning we can fit an essentially unlimited amount of data into the cache, but that data will quickly fade unless it is somehow physically transferred to the long-term storage neurons. This seems highly unlikely, if not obviously impossible. What kind of specialized neurons would be required to compose a part of the brain that has real-world physical limitations (which is part of what *defines* a cache) yet unlimited data capacity? Magic neurons would be required, and we don't believe in magic neurons.

Which leaves us with option two: the cache is limited both temporally (data fades) and in data capacity (something reflected by that item limit). But this option still has that flaw when considering our scenario: it offers no plausible way for the memory-wizard's narrative technique to enhance the short-term cache's apparent data capacity.

If instead, as Narrative Complexity hypothesizes, potential memories are laid into the same system as our long-term storage, our vast data storage banks and powerful associative capabilities in essence provide that unlimited data capacity (by linking to always-available open memory modules and/or creating links to existing modules). This model requires only the temporal limitation (represented by our half-life) & individual module data capacity (*our* item limit) to help create those unique effects of a short-term memory cache. And this model still allows narrative strategies to help circumvent these limitations when building something like a list. Within Narrative Complexity's system, there is an obvious way (previously explained) in which the memory-wizard's narrative technique can aid in overcoming a module's temporal limitations and data capacity.

In a short-term memory-cache model, there are only a couple of "logical" explanations for the effectiveness of the memory wizard's techniques. One, they're somehow skipping the short-term memory cache altogether and writing the data directly into his long-term memory. But this would mean that *all* narratively-structured data would have a chance to skip the short-term cache, which does not seem likely and would make a short-term cache much less useful (and almost arbitrary in its use, since all kinds of data can be arranged into narratives).

The other "logical" explanation is that his technique allows him to escape the cache's data capacity limits by linking the short-term data to long-term data outside the cache (a type of chunking). But this would not explain how or why narrative would help achieve this. In fact, using narrative to achieve this without clearing the cache would require a type of *infinite chunking*—the story allowing him to continue accumulating its narrative (& item-linked) chunks within short-term-to-long-term Russian-doll component parts (which is about as implausible as our magic neurons).

Consider that in a short-term cache memory model, even if he tried to use story to somehow link short-term data to pre-existing long-term data, he would still only theoretically be able to stuff the first 5–7 narrative chunks into the cache before encountering a storage problem. Thus (without employing *infinite chunking)* he'd still be forced to quickly transfer each half-dozen set of narrative chunks to long-term memory in order to clear the cache for new incoming narrative items that must occupy the short-term cache. (Which defeats the whole purpose of using memory devices like creating a story, since this explanation offers no reason why *all* kinds of lists couldn't also magically make use of this *way more robust* "just-transfer-it-to-long-term-&-keep-going" method simply by *deciding* to transfer that short-term data to long-term.)

None of these short-term cache mechanisms are very efficient or make much sense, and none take much advantage of all of the other mechanisms that appear to be simultaneously working to generate our consciousness. In the end, no version of any cache-based short-term memory system is very *elegant*. In contrast, Narrative Complexity's half-lives, modular data components, and narrative-building (discussed next) effortlessly-yet-interdependently create all of the apparent effects of a short-term memory cache—and all with the kind of simple beauty that has *the elegance of the human brain* written all over it.

The Architect in You

Another conundrum of cognition that a type of short-term memory cache intends to address: the notion of *working* memory. When we're thinking about something, our cognitive & computational tools need to know what data they're handling at the moment. Working-memory models typically suggest that its temporary cache of data is the reservoir for the information that our cognitive tools are currently using. Narrative Complexity views this cognitive processing—*linear processing*, which seems to occur in most people's left hemisphere—as a type of narrative-building machine.

(As we noted in our first essay, referring to this as "linear" processing is misleading because it suggests an *algorithmic* "computer-like" processing that the brain does not truly employ. However—from our theory's view—compared to the highly-associative nature of our right hemisphere databank, the much more organized & sequential nature of that genuinely *parallel* left hemisphere process of cognitive rule application is different-enough that its *results* are usefully described & distinguished as a *linear* process.)

Throughout these essays, we've acknowledged that narratives are, at their core, prediction tools. And in essence, most mathematic equations are exactly the same thing: prediction tools. "*2 x 2 = 4*" is, at its core, a prediction tool that we can use when encountering 2 pairs of objects and want to successfully predict the total number of objects without actually counting. This mathematic equation is expressing the same kinds of predictive relationships as the story: "*If Jill pushes Jack, Jack falls.*" In the case of narratives, seeing the event actually happening (Jack falling) is equivalent to "counting" in the mathematic equation.

In other words, there are two ways to determine the result of something: watch it happen (count) or foresee what will happen by applying a predictive pattern that imports current data and processes it using a set of rules (mathematic or syntactic) that have been proven to yield reliable (essentially, repeatable) results. This commonality of purpose & mechanic is why our brain's narrative- & sentence-building machine can also be a pure computational machine.

Another way to think of it: this is likely why our left brain seems to govern activities as diverse as writing essays and parsing calculus. In both cases, the machine is doing the same thing—building narratives/predictions by applying known rules to current data—it just uses the results for a very diverse set of problems.

For this cognition process to work, it needs that pool of data to draw from when building its narratives or pattern predictions. Since we've already shown that any short-term memory cache would likely be a clunky, inefficient (and implausible) add-on to our systems of consciousness, the question here isn't whether or not such a cache could serve as this pool of data (it could, although in the same clunky & inefficient way it handles short-term memories). Rather, the real question is: can Narrative Complexity's systems handle this need without the addition of a clunky appendage like a *working* memory cache? Our answer: *absolutely*.

When a thought enters our subconscious processing—laying down that potential memory & setting-off those memory-pinging associations to recorded data—whatever current or previously-recorded data emerges from the process (just-laid-down and/or "pinged" data that possesses the strongest & most-fluid linguistic, symbolic, emotional, physical & sensory associative connections) is inhaled by our narrative-building machine. This machine's job is to quickly sort and make sense of this data (*discern a pattern*) in relationship to whatever narrative, environmental, or physical problem/goal is on the table. (Our emergent linguistic data is accompanied by that emergent environmental/sensory & physical data—all of which is used to build these narratives.)

How exactly does our system determine this problem/goal, which is a necessary point of reference for narrative construction? This is, for me, one of those particularly hard problems of consciousness. Keep in mind that the "emergent data" that comes out of our subconscious processing has patterns within it, but in many cases (when it's not straight-word-for-word recall of one specific narrative parcel) it no longer has any syntax. Thus, it seems that it would be hard to convey the "meaning" of a problem, or provide something that could actually help direct intent (essentially, determine which rules are contextually appropriate here). So something else is probably happening along the way—*but what?*

In the view of Narrative Complexity, this process is likely aided by either our ever-busy switchboard, the thalamus, or by our corpus callosum, which connects & transmits information between the right & left hemisphere of our cerebral cortex. To understand how this "point of reference" is handled, it helps to look more closely at how our brain achieves these "loops" of data. Although we've been talking about these loops (& will continue to) as something akin to fluid data paths (like a race track that our horses of thought speed around) the process is more like a daisy chain.

As each specific neural network is activated, it nearly-simultaneously activates other networks—allowing pattern data to be communicated between these networks & "re-interpreted" or analyzed by the just-activated network, then sent along in its newly-configured form to other neural networks. This multi-faceted, constantly diverging & merging, looping data-relay occurs so quickly that if we were to watch it running with the naked eye, all of its various links would seem to be consistently lit. But in our brain, there is a very specific (yet dynamic) sequence in which all of this data moves from link to link, allowing that fluid & cohesive experience of consciousness to emerge (in other words, unlike poorly-dubbed foreign-language films, in real life people's words match the movement of their lips).

What does this have to do with determining that problem/goal necessary to build a narrative response? This daisy chain sequence of neural activations gives us the parallel loops necessary for our emergent data to be analyzed according to our problem. According to our theory, when a just-heard narrative parcel arrives in the right hemisphere of our cerebral cortex, the activation of these new memory modules likely triggers the activation of at least three other primary networks: the hippocampus (which helps sear the memory, something we'll explore later), the right-hemisphere network of stored memories (that data-pinging Google search), and the thalamus or corpus callosum (which aids in *narratively-contextual rule application*).

It's also likely that in the micro-moment before these three networks are activated (simultaneous to the memory modules being *initially* activated, not *in response* to their activation) the thalamus & basal ganglia attach current sensory data (which they've just received) to that memory. Then (as we just noted) *in response* to these modules' activation, the thalamus or corpus callosum helps to "translate" & pass along that just-recorded syntactic narrative parcel for use by our left hemisphere's narrative-building tools (to help employ that *narratively-contextual rule application*).

This essentially feeds our syntactically-defined "problem" (the previous thought) into our rule-based, narrative-building left-brain network as (or just prior to) that same network is also being activated (via the corpus callosum) by the emergent right-brain data that was just pinged (when that thought was recorded). In other words, you can use the syntactically-translated intent or need expressed by saying to yourself (or someone/something else saying to you) dialogue like *How do I get to John's house? (I must...)* or *My daughters are hungry (therefore...)* or *That part fits perfectly (thus...)* to aid in defining your goal or problem—which helps you to filter out & select which rules to apply when using that emergent data to build the most appropriate/useful narrative response or problem solution (which is, at last, a description of the actual process that we define as *narratively-contextual rule application*).

And when we look more closely at our daily lives, it becomes apparent that we frequently help to spur along even many of our *most mundane* (& seemingly rote) cognitive tasks by internally asking ourselves tiny narrative-prompting questions: *what am I looking for? where did I put that? where did this come from? why is this here? how did this happen? how do I do this? who the hell did* that? *etc., etc., etc.*

When the previous thought does *not* directly contribute to, trigger or help define the subsequent thought (because a "train of thought" has been interrupted or superseded by some other *higher priority* event or stimuli) our narrative-building systems can apply its most-foundational observational/causal syntax process to this emergent data pattern (a kind of pre-language mammalian cognitive syntax we'll describe later). This means the very basic environmentally-spurred thought (*The branch is falling! These children are crying!*) can be constructed from the emergent data *without* needing narratively-contextual rule-application, because this fundamental rule application is defined by that environmental (& primarily physical or spatial) context. Once this kind of simple thought kernel is fed into our machine, more elaborate narratively-contextual thought-extrapolation can begin.

Thus, Narrative Complexity hypothesizes that our cognitive processes build all this emergent data into dynamic narrative responses by applying all those different types of mathematic or syntactic rules to that emergent memory, environmental & physical data. *This* is that pool of data required for cognition ("working" memory). Our cognition processes sort & slot the pieces into their appropriate locations in the prediction pattern according to how each piece is defined (a word's meaning/ semantic content) and how each piece needs to be used (a word's function/ syntactic role).

There's one piece of clinical evidence that seems to contradict the mechanisms of this cognitive model, but that I believe actually speaks to the human brain's amazing flexibility & plasticity: cases in which individuals have had their corpus callosum surgically severed (typically in order to reduce debilitating epileptic seizures).

Despite removing this direct line of communication between the left & right hemispheres of the cortex, these patients remain generally cognitively capable [10] (although they usually display a variety of unusual, smaller deficits in perception & cognition). How is this possible if the corpus callosum plays a vital role in feeding emergent right-brain data into our left-brain's narrative-building mechanisms? In the view of our theory, severing these connections is indeed like removing the main data highway between these mechanisms—nonetheless, this persistent emergent data can still find "detours" around the new roadblock via other neural roadways (the most likely replacement route probably runs through that highly-connected & efficient thalamus, which is already communicating lots of varied data between these two hemispheres). In addition, this "re-routing" is also likely what causes some of those unusual deficits in these patients.

The exact ways in which the brain makes all of this happen—how it accurately matches so many different kinds of patterns to so many possible rules, how it re-routes data around roadblocks to access those rules—are some of the more deeply *unfathomable* aspects of the human mind. These "rules" are seemingly as plentiful & diverse as the memories themselves. It is difficult to comprehend how the human brain could create a system in which we are instantaneously & dynamically able to select & apply these rules to data emerging from our subconscious in such a fluid & successful fashion—even after a *severe* disruption of the system. And yet, think about how fluid your thoughts are, how quickly you take the words you hear from someone else, comprehend all their nuance & data, then construct a complex immediate response by applying these rules to your own emerging data. We do it effortlessly, and thus, we know that the brain can manage a system of such complexity because…*it does*.

As hard as it is to imagine such an unfathomably complicated system resulting from just a fundamental set of repeating, interweaving mechanisms—I believe this is what our brain is likely doing in the process of cognition. When our mind generates thoughts & solves problems using the most-recently-consumed or emergent, related & highest-priority data, it does so by applying a vast, diverse set of rules that help to create the presently most relevant or useful narrative, predictive or problem-solving equation.

Although this view of cognition helps to explain how the minds of exquisite scientific problem-solvers like Albert Einstein have used complex (& essentially mathematically-syntactic) equations to unravel & demonstrate their intricate, innovative solutions to a wide-ranging array of mysteries—this explanation doesn't seem to address how someone like Einstein *also* used his extraordinary cognitive *spatial* capacities to achieve the initial insights that led to those innovative solutions. In the view of our theory, however, our mind's use of its spatial capacities *requires* those word-based thought parcels in order to generate & manipulate the objects that we imagine.

In other words, Einstein still needed to *describe to himself* his thought experiments about things like riding a beam of light through the cosmos or 2 differently-placed observers witnessing the same bolt of lightning—and his descriptions helped him to generate & manipulate the visual data that he imagined within his mind, allowing Einstein to derive his insights by observing & analyzing his own complex conjurations. His extraordinary spatial capacities (suggested by some morbid *post-mortem* examinations of his brain) essentially made him capable of generating (& sustaining) more complicated, multifaceted, intricately-interactive visual scenarios from those word-based descriptions. Old Albert is proof that when someone can generate & manipulate both complex narrative/mathematic equations and complex spatially-based visual conjurations (and possesses a mind not inured to old paradigms) they're a decent candidate for solving some pretty cool & difficult problems.

Narrative Complexity's view of this complexly "inter-causal," multi-rule-based, syntactic narrative-building process is reflected in the theories of grammar presented by linguists M.A.K. Halliday & Christian M.I.M. Matthiessen in their book *Construing Experience Through Meaning: A Language-Based Approach to Cognition.* [11] Their work (which presents brilliant, highly-complex explanations of the mechanisms & powers of grammar & language) strongly supports our theory's central hypothesis of a language-based cognition process.

And to clarify a specific bit of language that *I just used*: the term "inter-causal" syntax is intended to convey both the way that a previous syntactic unit (a narrative parcel) helps to define the construction of the next syntactic unit (the process we just described) *and* the way that individual words within those syntactic units can interactively *cause* the transformation of each other (impacting the words' specific functions & meanings within that syntax).

You Know It or You Don't

Anyone who's familiar with the currently most-accepted view of these cognitive phenomena (*Dual Process Theory*) might immediately have a question here: where are the *two systems*? In the view of Dual Process Theory, the brain employs two cognitive processing systems (or types of "reasoning") that help us to respond to our environment: an implicit, subconscious, associative system (system 1) and an explicit, conscious, analytic system (system 2).

In many ways, these two systems roughly correlate to Narrative Complexity's primarily associative data-storage system and its primarily syntactic narrative-building system. The main difference is that Dual Process Theory posits that these two systems can work essentially independently from each other. In fact, the theory suggests that our associative system 1 can be used to make a decision or calculate simple tasks without even engaging system 2. In *Narrative Complexity* (similar to its handling of short term memory) *both* of these kinds of reasoning are handled by the mechanisms of our singular internal dialogue loop.

(Ironically, in this arena we find ourselves *disputing* some of the ideas supported by the author of our beloved *Prospect Theory*, Daniel Kahneman, who is a strong proponent of Dual Process Theory. Hopefully our ensuing explanation will justify contradicting the *patron saint* of our emotions theory.)

Without taking the time to explain Dual Process Theory in detail, the best way to show how our model handles these same tasks with greater simplicity is to explore some examples that are commonly used to explain system 1 & system 2. In Dual Process Theory, the problem "*2 x 2=?*" is supposedly handled by system 1. Here the system's associative mechanics provide the instant answer, "*4*." In this case, the claim is that system 2 has not been engaged because no real analysis or calculation is necessary. System 2 is engaged by the problem "*17x24=?*"—which does not provide an instant answer, but requires "conscious effort."

This effort is taken by Dual Process Theory as a sign that system 2 has been engaged—the calculation is explicit and demands our "analytical" processing. The theory often points to pupil dilation as a sign of system 2's engagement—something that I believe is merely a sign of more focused attention on a task, not a sign of specific systems being engaged. Our pupils, after all, also dilate during strong sexual arousal—which is not a case of anything particular analytical going on, but certainly a case of more focused attention on a task.

Using the mechanisms of Narrative Complexity these same exact system 1 & system 2 effects can be achieved simply by engaging our loop in different ways to solve different kinds of problems. The first problem (*2x2*) is merely an already stored (and very, very well imprinted) piece of data. When the problem enters our subconscious, the memory-stored answer of 4 pops right out in our emerging pool of data. And the narrative construction required to express this answer is *almost* non-existent. (Which is *not* the same thing as system 2 being unengaged.)

In fact, the ultra-simple response syntax of "*The answer is 4*" can be reduced to "*It's 4*" or even just "*4*"—because our brain isn't *stupid*, and it knows that in this case the only *truly* important syntactic element here is the actual solution. In other words, your brain hears the problem, 4 pops into our narrative building-machine, it drops the rest of the syntax because it's deemed unnecessary, and you shout out "Four!" before you even realize you're forming the words.

And because the answer arrives in your narrative-building machine with a *highly-valid* tag (and the problem itself isn't deemed highly important) there's no hesitation in responding or desire to actually apply a specific predictive rule to recheck our work through a true "calculation. " However, despite this answer's absolute obviousness to us, if our life *literally* depended on the solution, we might actually check that immediate pure-memory based answer by taking a moment to "calculate" (or maybe even ignore our predictive rules and count it on our fingers—I mean, our life *literally* depends on this here).

This is also why we're prone to be fooled by "trick" math problems that are essentially syntactically designed to fool us into arriving at the wrong answer. Dual Process adherents like to use these kinds of problems to demonstrate how system 1's associative reasoning is sometimes "flawed." But in the view of Narrative Complexity, when encountering these trick problems, people simply know that nothing particularly important is riding on the answer. And although narrative-building (system 2) is often referred to by Dual Process believers as "lazy" (which is why it "allows" system 1 to provide the wrong answer to the trick problem) I believe our narrative-building is really just trying to be as efficient as possible. This means that if an unimportant problem *looks* like (at first glance) a candidate for an obvious, high-priority rule or very-familiar stored data, we'll apply that rule or data, not check the answer, and move on. *That seems easy enough & not of any real significance, so we'll just make this assumption & move on.*

In response to these trick problems, it usually turns out that our assumption was wrong & we've been fooled by the "optical-illusion" syntax (in these problems, there is *always* a way to change the syntax in such a fashion that most people will apply the correct rules & get the problem right). *But who really cares that our assumption was wrong?* If our life was actually on the line, it's *very unlikely* our brain would accept the wrong answer without checking. Thus, in most cases when these kinds of errors occur —*it simply doesn't matter*. Which means our brain actually didn't make a particularly bad decision in assuming—after all, its job isn't to get *every* problem correct *all* of the time. Its job is to focus its highest resource-use on our highest priorities, and move through the rest as efficiently as possible. Meaningless mistakes are usually just that: *meaningless*.

Of course, we also make lots & lots & lots & lots (I could go on) of mistakes in rule application that *do matter*. But these cases are nothing like the intended-to-fool math problem. These meaningful mistakes aren't usually a case of "lazy" or trying-too-hard-to-be-efficient narrative-building—it's just a case, frankly, of *incompetence*. Bad rules, bad beliefs, bad rule-application & associations, a lack of useful memory data (I could go on). In other words—if it was important—you were probably *trying*, but trying just wasn't good enough. Our brains are awesome, but the humans that employ them aren't perfect.

Going back to our more straight-forward problem (*2x2=4*)—our response here is basically a super-quick, super-simple version of the loop. Nonetheless, there's no system or part of our loop that went unengaged—our memory was simply more relied-upon than our narrative-building for the answer. As with everything in our model, the whole loop *always* has to be completed for an actual thought or verbal response to emerge from us.

In the case of the second problem (*17x24*) it's unlikely that you've done that problem enough times to have a strongly-imprinted memory of the answer in your data-banks. Thus, after hearing this problem, your loop might first take a round or two processing internal dialogue about whether or not the problem is actually worth doing (*That's hard, but I get the point, I don't need to do this*).

If you choose to do the problem, you're likely to start applying some rules that help you to, essentially, syntactically divide it into parts that you've learned to calculate efficiently—something that might look like: *Okay, that's* (17 x 10) x 2 *plus* 17 x 4, so...*[(17 x 10 =170) x 2 =340] + [17 x 4 =68] =408*. And when you do this problem in your head, you're actually internally speaking those words to yourself—likely saying that first bracketed section as one narrative parcel, and laying *340* into your memory, so that data can be pinged in a moment and added to the *self-spokenly-arrived-at 68*. Again, like the simpler problem, this calculation relied on both our associative and narrative-building mechanisms to arrive at the answer—it simply relied on one more heavily because of the nature of the problem's difficulty and our *familiarity* with it.

(This "chunking" of cognitive tasks or calculations into more-easily managed components is an *intellectual* descendent of the *physical* process of *Motor Task Chunking*, which we'll discuss in our next essay.)

Once a problem becomes *familiar* (well-remembered) your brain can use its memory of the answer to provide that quicker response. So if you keep telling yourself right now "*17x24=408*" then tomorrow if you see that same problem, *408* will likely pop out almost as easily as *2x2=4*— even though *17x24* a much more difficult problem. But no magic of the mind has occurred here. The answer *408* is simply like any other memory data that's recent and has been repeatedly-recalled. And after time, if you never do that problem again, the data will likely fade and the problem will require more trips through our loop & more sophisticated rule-application to arrive at a solution.

And those essentially automatic responses happen in reaction to *all kinds* of incoming data. But, as shown, the instant response is not the result of our associative systems working *independently* of our more deliberate cognitive mechanisms. Rather, these responses simply require much less effort on the part of those narrative-building mechanisms, because the most-likely useful response has already been pinged in our databanks thanks to a previously identical (or nearly-identical) remembered experience.

Thus, when you see a vase start to tip, you automatically reach for it because you've seen a million things fall this way and your brain doesn't need to apply any rules to predict the result reliably. You see it tip, (your brain quickly shouts something like "Tipping!" or "That's falling!" or "The vase!") and you reach. But if the bookend tips over and starts a long chain reaction down the shelf that eventually knocks off your pencil holder (and that's never happened before) your first response to seeing the bookend tip might be to reach for the bookend and not get ready to catch the pencil holder.

Or you might pause for a moment, instantaneously scan the whole shelf, allow your brain to apply some physical rules to the scenario, and quickly (but not automatically) realize that you should be running for the other end of the shelf. Meanwhile, you're probably very quickly saying to yourself something like, "*That's tipping all those books…the pencil holder!*" or if you're really sharp, maybe just "*Bookend… books…pencil holder!*"

There are other implications that arise from the differences between Dual Process Theory's & Narrative Complexity's views of these systems (such as the apparent biases that Dual Process attributes to its systems) but exploring these differences isn't necessary to understand our theory's mechanisms. For now, the most important take-away here is that in *our* model, these narrative-building mechanisms that we're discussing are always at *some* level part of our conscious "reasoning" process, and *must* be engaged for any thought to emerge from our loop.

The Architect's Rule Book

Returning to our model's inner-architect and their syntactic, narrative-building rules —the next obvious question: where do those rules come from? There's at least one clear source of our rules: we learn them. In the view of Narrative Complexity, it seems absurd to assume that human babies enter the world with an understanding of *all* the myriad syntactic rules that govern sentence-building. Similarly, the narrative or causal rules that govern a specific skill-set (from chair-building to exploring physics) need to be learned through experience or study.

The other likely source of these rules at first seems more vexing to consider: we're actually born with them. This is vexing because it begs the questions: *What are these rules? What would they govern? How could they be purely fundamental & yet useful enough to begin building a complex, larger, inter-causal grammar?* According to our theory, these rules are the broader frameworks and most foundational principles—the type that help us to determine & recognize, for example, what a rule actually *is*, and how to construct new ones from the world around us. (Thus, *all* rules are ultimately built upon or somehow derived from these inborn rules.)

These are the kinds of rules that help us to understand—even before we've acquired language—that data usually requires a beginning, middle & end to make it useable. Thus, our likely-inborn fundamental rules are the rudimentary beginnings of *syntax*, whose first & most-basic purpose is to allow data to have start- & end-points—to define its limits & give it handles or borders, which are necessary to manage information as narrative parcels (essentially, as modularly-constructed but self-contained data packets).

And lest there be any confusion among adherents of "Universal Grammar" theories, what I am suggesting here is a much more scaled-down & fundamental-building-blocks version of inborn syntactic rules. ("Universal Grammar" theories propose that a broad range of specific & highly-sophisticated syntactic or grammatical rules have evolved to be inborn & essentially language-ready in all humans—a theory that's resoundingly debunked in Terrence Deacon's *The Symbolic Species*. [12])

How could a very young human brain's experiential recording mechanisms define such narrative or sequential beginnings & ends without the benefit of already-accumulated, rule-building life experience or without using the tools of language to "measure" such narratives? In essence, this is similar to asking: how did *any* pre-language mammal determine what defined a behavior- & prediction-aiding experiential data pattern as a self-contained, yet modularly-constructed unit?

More *specifically*: how did those earlier mammal brains (like dogs & monkeys) create non-linguistic-but-still-modularly-composed "proto-narratives" that allowed the determination of causal relationships and provided the capacity to use widely-varied, multi-sensory cortex-recorded experiences to aid in determining future behavior that helps to repeat (or dynamically create usefully-novel versions of) those causal sequences?

In the view of Narrative Complexity, our old friends *pain & pleasure* play a key role in catalyzing this process. Whenever strong pain or pleasure are experienced (i.e., an injury or a yummy) by pre-language mammals or very young humans, this experiential data module is viewed as a potential "end-point" (basically, as a gain achieved or a loss inflicted). Determining the "starting-point" of this narrative might then be as simple as identifying the most temporally-recent & recorded "high-attention" stimulus—a loud sound, a sudden movement, a novel scent, etc. (*basically, "spike" events that exhibit a certain category of specific attributes that allow them to be rudimentarily catalogued & cross-referenced as proto-narrative components*).

This kind of *retroactive narrative construction* is neurally possible because of the mechanics of "short-term memory" (or more accurately, the mechanics of priority-based data imprinting & the resulting memories' varying imprint "half-lives"). Those mechanics likely allow higher-attention/impact stimuli to hang around a little longer for soon-after pinging & comparison. In fact, this method of narrative construction might've been a powerful driver in determining how long a piece of recent data remains "viable" for possible use and thus, remains available to achieve longer-term imprinting. If recent experiential data does not attach itself to one of those pain/pleasure-spurred & retroactive narrative structures, the data is allowed to fade away.

Once these sequential, temporally-based end & start boundaries have been defined, it seems it would be easy to include other types of high-attention/impact *(spike)* experiential data (temporally-located within those boundaries) as different kinds of specific predictive modular elements within this narrative: high-attention/impact actions, reactions, events & "objects" (*inanimate* & *living*) that might be identified as (assumed) elements of causality within this sequence. This allows all these different elements to be rudimentarily categorized as proto-narrative syntactic components.

Of course, in a primitive system like this, there's lots of room for narrative-building errors, unreliability of data, and confusion between actual causes & mere correlations —which is why your dog's brain might mistakenly *assume* that spinning around 3 times before you fill its bowl is a *necessary* element of causality in the feeding sequence. It's also likely why the mechanisms of repeated recall (and its uses in strengthening frequently-pinged recorded data patterns) are central to mammalian cortex-based memory systems (recall uses that are *not* central to those earliest reptilian pain/fear-based amygdala memory systems). Those mechanisms were useful to mammals because they helped to reinforce memories of experiences that *repeated themselves*. This repetition essentially served as evidence of the original memory's *causal accuracy* (because the elements of a current narrative matched & recalled the original narrative, and generated the same result). This allowed repeatable (thus, presumably *reliable*) narratives to be more strongly remembered (leveraging their influence on behavior).

For this kind of mammalian proto-narrative, component-based & dynamic cognitive system to work efficiently, it would likely have to operate as a more primitive version of the same thalamocortical loop that's at the heart of human consciousness. Consider: in order to easily "go back a few steps" in one's experience & accurately temporally locate the likely "beginning" of a just-completed sequence, new incoming data must be sequentially fed into the same system that just recorded the data from earlier in the experience. As in humans, the experiential data loop in these creatures is like an ever-circling train that picks up new cars via sensory data input and drops them off in the brain's subconscious recording/associating mechanisms (where they hang around just long enough to determine if they were ultimately part of anything useful & worth remembering long-term).

This primitive system's effectiveness in generating useful, dynamic behavioral responses based on comparatively-related, cortex-recorded & narratively-constructed high-impact experiential data was likely a key driver in the development of the modern mammalian loop of consciousness. And if we shift our "wayback machine" into overdrive & travel into *hyper-speculation space*, we might glimpse the creature that I believe represents the earliest key evolutionary moment in the brain's journey toward this modern loop of consciousness: *lampreys* (jawless fish who were among the very earliest vertebrates—preceding sharks & *jawed*-fish).

Recent research on lamprey brain circuitry has revealed data pathways that I believe present a fascinating primitive correlation to our human loop: the integration of electro-sensory data (used to detect & track nearby movement) with visual data in the optic tectum via the dorsal thalamus (which will later contribute heavily to the development of the modern thalamus) & medial pallium (which will later contribute heavily to the development of the modern hippocampus, a crucial neural tool that we'll discuss next). [13, 14]

In the view of our theory, this is essentially the first appearance of what will become the thalamocortical loop of human consciousness. In addition to this circuitry primitively mimicking our own primary experiential data pathway, it also accomplishes something rather sophisticated: internally depicting (& tracking objects within) a multi-dimensional external environment via the integration of multiple sensory input sources (each of which are handling different kinds of stimuli in different ways, yet must "cooperatively" depict an integrated representation—a representation that critical behavioral & action decisions are entirely reliant upon). This is, essentially, the *very first* known appearance of that conscious-experience-inducing *internal model* that all vertebrate consciousness is built upon. In other words, *once upon a time* as lampreys swam about in those vast ancient seas, their sleek little selves were showing off a really, really cool new & super-clever way to view, interact with, and *experience* the planet earth & its creatures—a way that would hang around for a very, very, very (*and still counting*) long time.

~

Returning to us humans & that matter of determining narrative limits or borders in order to *define* an actual modular memory structure—this task leads us to a specific part of the brain: the hippocampus. Research has shown that in humans the hippocampus is primarily involved in both spatial tasks (area maps) and memory tasks that help create long-term memories.[15] The "modern" hippocampus (having slowly evolved out of that medial pallium) essentially first appears in amphibians, where it is only involved in those spatial tasks.

It makes sense that the first vertebrates to explore land needed a new spatial tool & system to help them navigate this new *non-fluid* environment. And the hippocampus conveniently appeared between those ancient creatures' now-expanding optical lobes and their age-old cerebellum—a perfect place for coordinating what a creature sees & maps with its locomotion.

It's not until early mammals that the hippocampus also becomes involved in the formation of memories—which (according to our theory) is also the same time that those modular neural structures begin appearing in those early cerebral cortexes. Thus, it's not hard to imagine that the hippocampus' original role as a definer of borders & mapper of space led it to take on a similar role in this new & suddenly very active process: the recording of modular cortex-based memories (and the hippocampus was *already* talking to the entorhinal cortex in the management of those spatial maps). [16]

Science has shown that the hippocampus helps to transform current or recent incoming data into long-term memory data, and damage to the hippocampus can cause problems like the inability to form new memories.[17] This would make sense if, indeed, the hippocampus is involved in outlining incoming data & defining it as a distinct narrative parcel—basically firing (and thus searing) a narratively-defined set of neurons together and creating one of those modularly-constructed but still self-contained memory parcels. If the hippocampus isn't working, incoming data essentially remains "undefined" in our memories; even if it is narratively-constructed, it's like an unending sentence whose yarn is always lost because it rolls perpetually away without ever being clipped & saved.

And in the view of Narrative Complexity, the hippocampus applies some of its own *inborn rules* (the kind that define those syntactic boundaries) when determining how that incoming data is snipped & stored—helping to create "modular memory maps" by employing some of the same tools that the hippocampus originally used to create its spatial maps. (And if you're looking for a neural model for how our hippocampus interacts with those right hemisphere memories, I'd explore the very recent discoveries about how a our hippocampus works with grid cells to create & maintain those detailed spatial maps. [18])

Another major example of an inborn or pre-programmed rule set is something we discussed at the end of our second essay: music. As we hypothesized, music seems to be a kind of pattern primer that gives our mostly-blank brains a set of basic data-relationship rules to model subsequent data rules upon. And the complexity of both the patterns of music itself & our emotional responses suggests that our brain could easily come pre-programmed with a full set of fundamental, but robust rules that our cognitive processes use as a kind of narrative-building starter kit and guide the dynamic creation of new rules.

Which brings us back to that other source —learned rules. How does our brain actually *create* new rules? When contemplating the creation of new rules, it helps to compare them with another predictive cognitive device—one that we explored in our emotions essay: *beliefs.* In the view of Narrative Complexity, the evolutionary roots of our belief-building system (likely spurred by learning to prefer cooked meat over raw) are actually found in this more-ancient rule-building process.

In our emotions essay we described beliefs as essentially high-value, high-validity prediction tropes. These beliefs are intended to reliably predict (across a wide variety of settings & circumstances) what will likely result from a specific kind of action or behavior. And these beliefs are *all* arrived at through study or experience (*no inborn rules here*). When we learn a belief-building pattern-prediction from a well-trusted source or if we have repeatedly experienced events (especially high-impact ones) that we perceive to support the pattern prediction, then it can rise to the level of a belief—leading the prediction to be more frequently & broadly applied.

All of these things are essentially true about syntactic rule-building as well. Rules are intended to reliably predict (across a wide variety of settings & circumstances) what will likely result from the specific usage of a narrative or linguistic syntactic element. In addition, we can learn a rule from a trusted source & immediately begin applying it (*a teacher explaining a rule of grammar*). Or we can learn a syntactic rule via repeated experience, which is appears to be the primary and by-far most effective method of rule-building. *(We learn best not just by being told what to do, but by subsequently doing it ourselves—preferably repeatedly.)*

Our brain is *trained* to pick-up on & build these kinds of rules through repeated exposure, experience & application. And like beliefs, all of this powerfully convincing (trusted-source or high-impact: "I'll never do *that* again") or repeated evidence helps to make a rule "stronger"— more likely to be frequently & broadly applied. *In pluralization, adding an "s" is essentially a stronger (higher priority) rule than unique pluralization. Thus, in any ambiguous or unfamiliar linguistic circumstance requiring pluralization, we will likely choose to add an "s" instead of attempting a unique pluralization.*

This experientially-based, immersive-learning process is the foundation of human language acquisition. And at the center of language acquisition is the construction of another key narrative-building resource: our *vocabulary*. Science has speculated that our brain contains, essentially, a dictionary of words that it builds over a lifetime. [19] In our theory, this vocabulary resource is distinct from the words stored in our memories, although those memory-stored words are the original source of (and continually help revise) this dictionary. Just as we build distinct rules & beliefs from the patterns in the emergent right-hemisphere data that sparks them, we build our vocabulary of words from the same pool of emergent data.

To describe those words in more specific systematic terms, in our view they are, essentially: modular cognitive/neural components that possess a wide array of defining semantic & functional attributes and external associations—all of which can vary in malleability & strength, and that together determine the full range of the word's meaning, syntactic capabilities & symbolic capacities (content that can be—but is not *always*—embellished or revised with every experienced or studied usage of the word).

And according to our theory, these words are, at their core, comprised of & represented by their *phonemic components*. This means that when those word-based parcels of just-heard internal dialogue arrive in our right hemisphere for recording, they arrive (& are seared in those memory modules) as collections of *sounds*. Thus, when we talk about *language-based* memory & cognition we are actually talking about *sound-based* memory & cognition. (In deaf individuals, these components are likely visual and/or gestural-physical.) Every word's semantic & functional attributes are attached to or built around a core of sounds (essentially, a phonemic neural footprint) that serves as a word's unique & specific "identifier" (you might think of this phonemic footprint as a word's *social security number*).

The letters that comprise a word are also, obviously, very closely tied to its phonemic footprint—but written language is, ultimately, an *external* tool and still built around (& merely makes *extraordinary* use of) that phonemically-based process of language-based cognition. Keep in mind that humans were using verbalized words & internal dialogue to employ these language-based systems of memory & cognition long before written words came into the picture, and thus, those externally & internally heard words *must* have been represented by neural components that were based upon the sounds that comprised the words (aka, a unique phonemic footprint).

Moving on from words themselves and returning to that larger left-brain *vocabulary resource*, I also think it's very possible that this word-based vocabulary resource actually contains another set of items: people. I believe that the names of people we know are stored in our vocabulary like words; those names contain our most fundamental “definition” of that person. (If not actually contained *within* this vocabulary resource, our “people resource” is still likely a similarly-constructed & closely-related resource that’s used at the same point in the narrative-construction loop.) And the most important part of that definition ultimately: whether or not the individual is judged as an *Agent of Gain* or an *Agent of Loss* (discussed at length in our emotions essay).

This would be the perfect place for our brain to store this person-associated value (an essential element of emotional production). And a mechanism like this would allow this *potential* value judgement about someone to remain separate from—while still being impacted by—a *known* gain/loss judgement about them, recorded in our data storage. The same kinds of processes that we use to convert emergent memory data into rules & other vocabulary could also be used to define people & calculate their current value. And this catalog of people (represented by their names) would be a resource as vital as words themselves when building these narrative parcels.

But proposing such a “people resource” leads to an important question: how would our brain determine when a word that’s used for a creature (or object) gets filed as a unique “name” in our people resource and when it’s just filed as a regular object-defining word in our *general* vocabulary resource? Within our model, there appear to be at least a couple of methods for making such a distinction. The more deliberate & less intuitive way is simply to *give* the thing (living or inanimate) a unique name and to *repeatedly use* that name when interacting with & referring to that specific thing. In other words—thanks to that looping one-thought-leads-to-another internal dialogue process—our cognitive systems simply respond to that initial self-designation of this word as this specific thing’s unique name (leading the named thing to be filed in that people resource in the next round of processing). However, because any object is also closely associated with its object-identifying word (in addition to any unique name you might’ve given it) the *repetition* of that name-usage is key to ensuring that our cognitive processes handle that thing as part of our people resource and not just as a regular word.

And this distinction is important because once we’ve placed this name in that special people resource we’re highly prone to begin treating that creature or object like (& to think of it as) an actual person—*even if that thing is really just an un-alive, inanimate object.* This is why we can often have seemingly-illogical, but clearly-felt emotions for named (thus, *personified*) & revered non-living objects like our cars or stuffed animals or *pet rocks*. Of course, in cases like those stuffed animals—and those pet rocks, if you’ve painted a little smiley face on it—our brain is also making use of that other *less deliberate & more intuitive* method for judging other things as "people" (or at least as things worth naming & storing in that special resource): our mirror neurons.

As we discussed in Essay #2's exploration of emotion's evolution, mammals have long been using mirror neurons to help identify emotional states in other "like entity" creatures (as demonstrated by those rats who used mirror neurons to reflect the stress experienced by other rats). This means that our brains are built to automatically identify & analyze the human-like features & actions of other creatures (& objects)—intuitively leading those things to be handled differently within our cognitive & emotional systems.

In other words, because a willow tree doesn't in any way resemble a human, when its leaves shiver in the wind we don't automatically (& essentially *empathically*) intuit that the tree is scared or cold (and thus, we don't experience any emotional response to this observed shivering). But if you've simply drawn a face on a rock (and haven't even *named* it yet) then drop a heavy object directly on top of that smiling rock, you're likely to have a tiny, momentary (yet almost unavoidable) empathic *wince* or internal ping of *ouch* in response to witnessing this *merely metaphorical* cruelty (and/or you might enjoy a sinful little giggle over the poor helpless rock's misfortune). In terms of our people resource, this suggests that—in addition to our more deliberate method of simply designating (& repeatedly using) a unique *anthropomorphizing* name for a specific thing—our mirror neuron-based method of identifying & analyzing other things as "like-entities" also plays a key role in helping to shape & determine the contents of that people resource.

Now, because we're addressing how individuals speculate about the inner feelings & experiences of *other* individuals (or rocks) we've wandered into the territory of something that's referred to as "Theory of Mind" (or ToM, as it's commonly abbreviated in academic literature). Basically, Theory of Mind is defined as an individual's capacity to understand (or tendency to perceive) that another person (or rock) *has their own mind*, and thus possesses their own unique intentions & feelings & inner experiences that all motivate (& possibly predict) how that specific person will behave (or feel) in response to some stimulus or event.

Throughout history, much of philosophy has treated ToM as a kind of special case—essentially, as a unique capacity that's tied to a specific brain "module" or type of cognitive process that enables this kind of "mindreading" (aka, internally modeling the internalized experiences of others). However, as with most of the more complicated aspects of cognition, our theory views this process as a result of *multiple* systems working (as they always do) *in concert* to handle the various cognitive tasks required for understanding (or guessing at) the feelings & intent of others. (For those familiar with the various philosophical approaches to ToM, our view generally fits within a *simulation theory* approach—and for those unfamiliar with that approach, consider this parenthetical a rude & unnecessary interruption.)

Ultimately, according to our theory, this ToM capacity is a result of how our cognitive narrative-building & emotional systems *make use* of that people resource and those specialized mirror neuron mechanisms. Think of it this way: when we try to apply this "ToM capacity" to perceive or determine the inner experiences of others, what are we really doing? *We're predicting how they feel*. Sound familiar? Yes, ToM is really just another version of our cognitive & emotional systems doing *what they're always doing*: predicting results and making decisions based upon those predictions.

Because those systems are so diverse & versatile, there are actually multiple ways for us to make predictions about other people's inner feelings, desires & intent (and to take action based upon those predictions). Some of those ways are more intuitive & automatic (thus more useful for quick decision-making, but more likely to produce incorrect predictions if multiple factors are involved) and some are more deliberative & complicated (thus less useful in quick decision-making, but less likely to produce incorrect predictions if multiple factors are involved). To demonstrate, let's look at an example…

Instead of buying it online, Bob goes to a real-life bookstore to get a gift for his uncle's birthday (hey, it could happen). *As he scans the table of new books, his eye catches the cover a World War II tome—even though Bob himself is totally bored by war books, he smiles widely and nods, then reaches out to grab the book.*

I'm betting *all of you* just did the same thing —*you assumed that Bob's uncle desires books about World War II*. But you don't even know Bob's uncle! What are you, some kind of *literal* mindreader? *You wish*. Although you likely came upon this conclusion about Bob's uncle's inner desires pretty quickly, you actually used a little bit of that deliberative cognitive process: predicting the uncle's desire by analyzing circumstantial narrative elements (like the stated purpose of Bob's purchase & his response to seeing the book).

And it's easy to mess with this kind of narrative *assumption* (as opposed to a genuine automatic *intuition* of someone else's inner experience, which we'll discuss in a moment) simply by changing a narrative element: *Bob really, really hates his uncle.* Now when you think about Bob's smile at seeing the book, you might *assume* (aka, *predict*) that Bob's *intent* is possibly more mischievous or malicious—thus altering our *assumption* about how his uncle feels about WW II books… *maybe Bob's uncle is a peace-freak who actually hates anything about war.* We seem to be pretty good at predicting the desires & intent of the previously unknown Bob & his uncle purely based on a few nuggets of circumstantial evidence.

And we're able to make these predictions—without actually *knowing* either of the men—simply because *we know they're both people*. Therefore, when we initially encounter (& internally process) the men in the story, we designate them as *people* within our vocabulary resource. This leads them to be treated within our cognitive narrative & emotional systems as one of those people-resource-stored *Agents* (i.e., of Gain or Loss) as we process the rest of the story. Essentially, our cognitive systems re-orient our narrative prediction-building & emotional gain/loss, etc. analysis from ourselves (& our own desires) to the perspective of the Agents in the story (Bob & his uncle). These systems then do *what they're always doing*—make predictions that are based on (& produced by) the specific narrative and our emotional analysis of that narrative.

Returning to Bob's own act of "mindreading," we can see that he's using a method that's different from what we just used. If we go inside Bob's head to examine the kinds of predictions he was making about his uncle's desires, we can see evidence of a more automatic & *intuitive* method for predicting how others will feel & respond—a method that we can use when we *actually know* that other person (or are standing right in front of them when something happens). This method is why Bob might almost *automatically* smile when he sees the World War II book (even though he personally dislikes them).

Because Bob knows that his uncle uniquely loves such things (and this attribute is attached to that people resource "definition" of his uncle) and because Bob's intent in that moment is to find something his uncle will enjoy, those two pieces of data employ the emotional analysis that produces that instant *this-is-perfect* smile—an analysis that also, in the subsequent moment, allows Bob to do another (but now slightly more narratively deliberate) "mindread" and predict his uncle's inner state of happiness upon receiving the book, all of which ultimately leads to Bob's book-seeking *action*.

Furthermore, if Bob is standing right in front of his uncle when he opens the gift and witnesses his excited response, Bob's prediction about his uncle's inner state of happiness in that moment is likely also being shaped by those mirror neuron systems that help Bob *intuitively* identify his uncle's emotional state and thus *perceive* his inner feeling (and respond appropriately based on that prediction). The flaw, however, in more quick-action oriented & intuitive "mindreadings" is revealed if we consider another slightly-altered version of our Bob-&-his-uncle scenario: *his uncle already has a world-beating collection of WWII books and just started collecting Civil War tomes in their stead.*

Now the fact that Bob has always defined his uncle as a WWII buff (combined with his automatic, intuitive *people-resource-definition based* method for predicting his uncle's inner desires & consequent excitement) might lead Bob's "mindreading" to result in a poor action choice & an undesired outcome. However, if Bob takes some time to specifically recall his last visit to his uncle's home, that might spark consideration of some of those other mitigating factors—allowing Bob to adjust his initial automatic, intuitive (& incorrect) prediction about his uncle's inner feelings (*he'd love a WWII book*) and to produce a more accurate prediction (*he desires a Civil War book and would be disappointed by a WW II book)* by using those more deliberative & complicated methods of analysis that are capable of taking into account multiple factors in making such a prediction.

In the end, if Bob & his uncle have taught us anything here, it's that "Theory of Mind" is not so much a *special case* or *specific capacity* as it is just another multifaceted & contextually-varied result of human brains *doing their thing*—and applying that thing (*prediction*) to the motives, intents & feelings of others.

Leaving Bob, his uncle & ToM behind, and returning again to that specific people resource & all those persons of interest filed therein… Who's the *most important* person in our lives & minds? *Numero uno:* ourselves.

Much of the latest research strongly suggests that *self-related* descriptive data (personal traits, abstract qualities, behavioral characteristics, symbolic individuals or items, etc. that we associate with & define our identity) is stored separately from all of that narratively-based, associative, right-brain data. [20] And if we're looking for a likely left-brain location for this *definition of self*—that dictionary containing the definitions of words & people seems like the perfect place to stash *us*.

These word-, people-, & self-filled vocabulary resources are likely assembled & applied in the same loop locale as rule-building. This is *because* of the role words play in assembling a narrative parcel. Many of the words required to complete a parcel's syntax likely come straight from (or are *direct* vocabulary matches from) the emergent pool of data. But this mathematic or linguistic syntax also requires other words, the connective words and/or words that need to represent previously unassigned "values" (essentially, numbers or ideas) that are also part of this new equation. Those other words are drawn from our vocabulary.

Think of it this way: when we watch Jill grow hostile toward Jack and anticipate her pushing him (and Jack falling) our brain basically says "If Jill pushes Jack, he will fall." Here the sights of Jill & Jack come from that pool of emergent data (providing a direct correlation to their names). But it is the observation of Jill's hostility (not the sight of her pushing him, which hasn't happened yet) that's the actual *source* of the word "push."

And where exactly does the word itself come from? That vocabulary resource. When Jill's hostility data emerges, it helps us recognize a pattern in the moment, which calls up related rules—which in turn define the linguistic syntax used to express this pattern. The data also helps us to choose an appropriate word from our vocabulary to represent this value or idea as required by the defined syntax. Basically, during this syntactically-based narrative-building process, our brain has three main sets of resources that it applies to emergent data: a set of inborn rules, an accumulated set of learned rules, and an accumulated set of learned vocabulary. (And right beside them on our neural shelf is that accumulated set of beliefs.)

Keep in mind that all of these resources (although probably more-neatly organized & prioritized) are still using the same basic kinds of neural structures that our data storage uses. Thus, each of these massive collections includes within it a wide array of associations between the different rules or words. Our efficiency in managing *and* our individually-unique handling of these rules and vocabulary is likely affected by the way in which we've set-up these associations between them. Great "thinkers" (scientists, writers, philosophers, professors, etc.) likely have very-efficiently arranged & prioritized sets of rules governing their area of speciality.

However, this kind of rule-system & linguistic efficiency is not likely the same as what we generally consider to be *intelligence* (which reflects neural abilities that are very difficult to improve). We'll explore intelligence in detail shortly, but here's a quick example of why this is true: when we take an IQ test (designed to specifically judge "intelligence") we aren't actually using our system of learned rules to discern & respond to patterns. Rather, we are recognizing & applying patterns that are intended to be demonstrated *within the question itself* (and IQ test answers intentionally do not require a deep vocabulary). Thus, these kinds of tests isolate our more fundamental (and likely inborn) pattern recognition & application abilities.

And the essential sameness between rule/vocabulary/belief-recognition/building/application appears to be another effect of our brain's looping elegance. All of these resources are assembled & applied at basically the same point in the loop. As soon as our brain builds a narrative using rules & vocabulary, it immediately judges it for necessary emotional production. Thus, beliefs are used to help emotionally-analyze a narrative in the adjacent micro-moment after rules & vocabulary are used to build the narrative. This means that very closely-related parts of the brain would likely handle these three (*extremely*) broadly-based tasks. (And at the end of this book, you can explore our own very *rudimentary* model depicting a hypothesized data pathway through those closely-related brain areas as these systems are engaged.)

Show Me! Show Me!

Fundamentally speaking, this whole rule/vocabulary/belief-building process uses the same simple technique that's at the root of how our brain builds all of its systems from the ground up in a mostly-blank mind. It is using accumulated correlation to help determine rules of causation. In short, to our brains, repetition equals truth.

There appear to be at least two main reasons why our brain is so well-suited to use this deceptively-simple, correlation-leads-to-causation mechanism when building its self-defining architecture. One reason, that amazing loop. Here's a cool view of the loop that we haven't shared yet: it's basically our brain's way to apply the scientific method to human experience. We begin by observing (our external & internal data input systems) then we analyze (that subconscious process of associating, comparing & evaluating data) then hypothesize (our narrative/prediction-pattern building) then test our new hypothesis (act, speak or behave as a result of the narrative/prediction) and finally observe that result, beginning the loop again.

Amusingly, this cognitive-analysis sequence also exactly matches an acronym that was taught to me *many many moons ago* by an unlikely *Agent of Gain*—Mr. Kurtz, my high school driver's ed instructor. The acronym: SIPDE—*Search Identify Predict Decide Execute* (which is still a sound driving strategy). The more familiar you become with the basic neural principles behind our experience of consciousness, the more you realize how frequently they seem to "accidentally" duplicate themselves within every aspect of culture.

(In fact, if you really want to *freak yourself out* & become suddenly *over-aware* of how deeply & powerfully words have engrained themselves into the way that humans interact with the world: next time you're in a retail establishment, take a good look around at all of the *words* that are plastered *everywhere*, addressing *everything*. Product content, use & category, store organization & procedures, "lifestyle" & marketing messages, special product & service enticements, legal disclaimers, employee rules, name tags, exits, *etc.*, *etc.*, *etc.* Even in our heavily image-based & visually-overstimulating modern world—a huge portion of that *overstimulation* in our urban settings comes in the form of huge volumes of everywhere-in-sight *words*.)

Just as the repeated application of our *not-so-coincidentally-brain-loop-based* scientific method has helped humans to build a set of rules that govern construction within our physical universe, our internal dialogue loop uses this same process to build our own individual set of rules that govern construction within our mind. And because this loop is perpetually running at unimaginable speeds, it's able to conduct an almost uncountable number of tiny, rule-building experiments over the course of a lifetime.

Which leads us to the second reason why our brain is so well-suited to use this deceptively-simple mechanism to build its self-defining architecture: that extraordinary *computational depth* of our mind. In order to effectively build, manage & apply this massive collection of rules, you need a machine like—well, a machine like the one-of-a-kind human brain. And when you have that kind of processing power at your disposal, a seemingly-simple method like correlation leading to rules of causation can still result in a creature of amazing (and amazingly accurate) complexity.

As we noted, resource-building occurs—conveniently & elegantly—in exactly the same loop locale as resource-application. How do the same mechanisms handle the tasks of both building & application? Think of it this way: when our brain seeks to *apply* rules (& words) to that emergent data, the first thing it must do is discern a *pattern* in that data—so it can determine which rules will be used in narrative construction (undertaken in relation/response to our problem or goal). Part of this pattern recognition is a matter of matching emergent data to those learned rules. Another part is a matter of matching that data against those more fundamental inborn rules that define syntax itself.

When new data contains a pattern that exactly matches a learned rule, it reinforces that rule—makes it *stronger*. This pushes it further along that spectrum of correlation becoming causation. The causation "threshold" (likely determined *on a curve* based on our current hierarchy) is essentially the point where a pattern's validity/reliability scores high enough to qualify it (in our flexible hierarchy) as a rule or belief.

When new data contains a pattern that doesn't match any learned rules, but still matches some of those fundamental rules (thus defining it as a usable syntactic pattern) then our rule system takes that new pattern & makes it a new rule. *This* is one of those ways in which we build our resource of rules. Unfortunately, in these cases—because this is a rule's first appearance in our hierarchy—it's likely very, very low on that rule totem pole. This makes it easy for the rule to go unapplied —even when it's useful. In fact, the just-born pattern is barely a real "rule" at all. But this is how the *source* of this new rule can help. If the source is well-trusted (or involves a high-impact event) then the data is immediately judged as highly valid or significant, giving it greater prestige (*aka, more-fluid-&-likely-to-be-accessed incoming pathways*) when this first-timer is placed in our rule hierarchy.

Nonetheless, a rule doesn't even have to come from a *consciously-known* source to be built. Our brains automatically ferret-out rule-building pattern data from *every* experience. Having a "teacher" is simply a case of someone calling a rule to your attention, allowing you to rapidly accelerate that immersive, *soak-it-in*, rule-learning process of experience.

There are also those cases when new data contains a pattern that *partly matches* an already-learned rule or the new pattern *contains within it* an already-learned rule—and *in addition* this new pattern is also judged overall as a fundamentally valid expression of syntax. These new patterns can also become new rules—ones that have essentially been *built upon* or are *variations* of a known rule.

Beliefs & vocabulary are built in essentially this same way. The primary difference between these beliefs, rules & vocabulary are their purposes: beliefs are used to influence our actions out in the world (decisions & behavior) rules are used to influence the actions within our brain (narrative construction & syntax) and vocabulary is used to create definitions in our brain. Ever wonder why we all seemed so obsessed with stuff like top ten lists, rankings, and "commandments" (regarding pretty much *anything*) or why we seem to prefer viewing everything in our world as some sort of hierarchy? You can stop wondering. We're addicted to this stuff because our brains can never really get enough rule-building data—our minds are rule-building & hierarchy *junkies*.

~

There's another aspect of linguistic expression that aids humans in adding meaning to these kinds of syntactic constructions: *inflection*. In the view of Narrative Complexity, inflection is essentially the result of applying emotional data derived from the *neurally-built* version of a narrative parcel to the subsequent *physical expression* of that narrative parcel. In other words—the narrative is built, it's analyzed for emotional production, then the narrative & its emotional data are sent (probably via the thalamus) to motor control areas that use this combined data to help determine *how* the language is expressed, *aka* vocal inflection.

Matching emotional expression to specific linguistic elements during the physical act of speech is a task that would seem to require more than just a purely motor-focused part of the brain—since the task is one of analyzing data (determining exactly *how* the emotion will impact the expression of certain linguistic elements) in addition to producing the necessary motor instructions. Thus, our theory hypothesizes that the key player in the matter of inflection is an area that research has suggested presents a unique combination of linguistically-analytical & motor-controlling capacities: Brocha's speech area (which appears to be the neural locus of speech production). In addition, it seems likely that Brocha's correlating "input" area —Wernicke's speech area, heavily involved in the analysis of heard speech—plays a similar role in interpreting the inflections in other people's speech.

Evidence of the separation between these tasks of syntax-building, emotional interpretation & inflection-application seems to be found in the kinds of *conflicted* inflections common among young toddlers. I've observed that young children (like my own) who are first developing their language skills will often express a conflicted or *uncertain* combination of inflection & language use. For example, saying "No" with upward-lilting inflection (essentially, an *uncertain* inflection—as opposed to the more natural downwardly inflected negative response) when the "No" response—even though grammatically correct—actually reflects consent.

Dad asked, "You never want to go anywhere without your blankie, do you?" The toddler replied, "No-ooo…" (with an ascending "ooo…"). Most adults in this situation would automatically "self-correct" this kind of expression and give the negative language its more common "downward-leaning" inflection—even though this linguistic response conflicts with the fact that they are technically expressing consent. *Dad said, "You never want to go anywhere without that iPhone, do you?" The teen replied, "No way, man."*

In the toddler's example, the detected-conflict's impact on the neural moment of inflection-application suggests that word choice & emotional analysis occur both separately *and* prior to the assignment of expression to each word choice. Thus, in the actual speech production, the word "No" was processed both according to the *functional/emotional* role it served in the syntactic structure (consent) and its separately-defined semantic content (rejection) —causing a conflict in the inflection that resulted in the *uncertain* (upward-lilting) expression of the "No."

Thus—because this process separately accounts for emotional & semantic content when determining inflection—a word can be inflected in a purely semantically-determined way, regardless of the word's functional/emotional role or vice versa or in some combination, depending on intent. In toddlers, their developing brain seems to have more difficulty in confidently resolving such inflectional conflicts, whereas adults seem more capable of flexibly adjusting inflection based on syntax, word-meaning & intent.

Because of the “musicality” of inflection & the innate use of basic inflection by toddlers, the rules that determine how speech is inflected are likely fundamentally inborn—and closely related to those inborn musical rules, which (as discussed in essay #2) specialize in analyzing & structuring data relationships according to emotions (the essence of inflection). This inborn capacity to apply & interpret inflection in *any* verbal utterance (even *before* an infant or toddler has developed a true capacity for language) likely helps us to construct (with the aid of those inborn *syntactic* rules) that *initial* basic neural framework necessary for developing the complex (& primarily *learned)* linguistic & cognitive processes that sustain human consciousness.

Of course, like almost all of our cognitive rules, these inborn foundational (& essentially musical) rules of inflection can all be revised & embellished according to experience—allowing for all of those individual (& cultural) tendencies of inflectional (& musical) expression. If our inflection mechanisms do, indeed, work in this fashion, then it suggests that (in addition to its other speech-producing duties) Brocha’s area helps to “couple” emotional data with semantic content in the production of inflected speech, and Wernicke’s area helps to “decouple” emotional data from semantic content in the interpretation of inflected speech.

In other words, this whole process of syntactic construction, analysis & expression is like a gigantic rule-, emotion- & belief-application festival. Each step along a thought’s path from from our subconscious to our lips involves another layer of hierarchical analysis & application, helping to determine everything from the words we say to how we say them *before they’re even spoken.*

The Great Syntactic Divide

Despite their cognitive similarities, the differences in purpose between all those rules, vocabulary & beliefs lead to an important distinction in how these resources appear to be built. The distinction: rules & vocabulary are built (& applied) *pre-syntactically*, and beliefs are built (& applied) *post-syntactically*. (And inflection, which is an ultimate motor result of this construction, is handled after both of those processes.)

This means that rules & vocabulary are built from (& applied to) the patterns identified in emergent subconscious data (which leads to the application of syntax to that data, thus "pre-syntactically") and beliefs are built from (& applied to) patterns detected in those syntactically-constructed thought-parcels (thus, "post-syntactically").

This essentially means that rules are based on "facts": pure data that can be arranged & matched to an identified valid narrative/prediction pattern, a pattern which is—or was at its root—derived from our inborn rules. In contrast, beliefs are based on our *interpretation* of those "facts"—in essence, what those facts *mean to us* (emotionally) according to the syntax in which they have been structured.

In other words, our beliefs (like *all* of our emotional mechanics) are behavioral guidance-&-prediction-patterns based on how we interpret the relationships within those syntactically-constructed "factual" data patterns. And these belief-defining behavioral patterns are *all* learned in some fashion or another over time (which is *different* from the rest of our emotional mechanics, whose behavior-influencing gain/loss equations & responses are all *inborn*—i.e., even our pre-toddler & belief-less selves *automatically* feel emotions like *sadness* when someone takes away our lollipop).

Now let me un-spin your head. First, here's a quick way to tell if your brain has constructed one of these high-level prediction patterns as a belief or a rule: *how do you feel when you you violate this belief or rule?* When we violate one of our beliefs, we feel *guilty*. When we violate a rule, we just feel...*stupid*. Thus, when we cheat (if we *believe* cheating is bad or wrong) we feel guilty. But when we violate a rule of grammar, we don't feel guilty, we feel incompetent.

And this doesn't just apply to rules of grammar. If we fail to apply reliable rules of narrative causality or physicality—leading to a bad result or an incorrect prediction—we aren't likely to feel guilty. Rather, we're likely to be dismayed or perplexed by our mistake, asking ourselves things like "*how did I not see that coming?*" In these cases, we don't feel that we *chose* our error; in fact, we probably *thought* we were applying our rules correctly at the time. Thus, the mistake merely makes us feel like a failure, not like a bad person. Violating a belief, of course, makes us feel exactly that way: *like a bad person*. In this case, we feel that we *did* choose our error (or felt powerless to resist its temptation) *despite* the fact that we knew what we were doing was "wrong" (likely to lead to an ultimately bad result).

This pre- & post-syntactic application of rules & beliefs likely plays a key role in the way that we consciously perceive these different kinds of "mistakes." Rules (pre-syntactic) are applied to emergent *unconscious* data—at the very beginning of the narrative-construction process—which is why overlooking or misapplying them feels like an unconscious mistake. But beliefs are applied to already-constructed parcels of dialogue (post-syntactically) and generate specific feelings (feelings that are *intended* to immediately call attention to any belief violation or compliance)—both of which (the dialogue & feelings) ultimately appear within our *Dynamic Core-based* arena of consciousness. Thus, we feel like we are *consciously aware* of our belief violations at the time we commit them, and are therefore *responsible* for our mistakes.

Which leads us to a probable truth that you aren't going to like—even though one of these actions (belief-application) feels more "voluntary" than the other, they're both essentially the result of very similar kinds of mechanics. Another way to look at it: our belief-application system (which is at the root of most of the big decisions that we *feel* we make voluntarily) is not any greater an "Agent of Self" than our rule-application system. They play equally vital & closely-related roles within the syntax-based systems that our consciousness uses to build predictions & make decisions. But, as noted, they're applied on opposite sides of syntactic construction within our internal dialogue loop.

Which is, come to think of it, actually a pretty *big* distinction—it's that *Great Syntactic Divide.* Is this distinction enough to say that our belief-application system is where the notion of "free will" might start to get a foothold in our consciousness? That's a delicious & dangerous question—and one that we'll save for our next (the *final*) essay. For now, it's more useful to focus on these systems' *similarities* in addition to their differences. The many similarities between beliefs & rules mean that we can often interchangeably use different combinations of both resources to arrive at or frame a decision. Examining an example of this should make everything here a lot more clear:

A high school student is taking a difficult test in a room full of other students. (The test is not graded on a curve, and no one powerfully admires the teacher—we'll note why these factors might be important later.) During the test, the teacher is called out of the room on an emergency. She says she'll be only be gone ten minutes, and that she trusts no one will cheat in her absence. She is, of course, wrong. As soon as she leaves, everyone except for our one student immediately begins using their notes and books. Our student hesitates, then finally thinks…

Now, our student could obviously think a plethora of things. But if they are going to eventually decide to use their notes or not (as opposed to the decision causing them to pass out from the pressure or run out of the room screaming) then their decision-making thought can likely be reduced to one of the following types of narrative constructions (essentially, types of belief- & ruled-based reasoning). I've labeled each example in order to help distinguish & define the different types of narrative constructions.

All of these constructions assume our student believes at some level that "Cheating is bad" & that everyone would benefit strongly from cheating (obviously, if they didn't think it was bad, they would simply cheat, and if there wasn't a benefit, they wouldn't have any need to cheat—except for a need, say, *to not look like a square*, which will also be covered). We'll start with the most-obvious construction…

Belief failure: *I'm using my notes, which I know is totally cheating & I don't feel good about it, but I want a better test score—end of story.* There's not much to explain here. The potential gain was simply too tempting for this person & their belief lost the decision-making neural war—which can happen for lots of reasons: weak beliefs, strong need, big potential gain, ingrained behavioral patterns, etc. This person is likely to feel a good dose of guilt (& it's probably a familiar feeling to them).

Rule-based rationalism: *I'm using my notes, because everyone else is too, so it's not even actually cheating—it's basically an "open book" test now.* This person has found a way to avoid engaging their "Cheating is bad" belief by constructing & defining the narrative such that the act does not constitute cheating. This person is likely to feel little (if any) guilt about the act.

Belief-based rationalism: *I'm using my notes, which yes, is technically cheating—but everyone else is doing it.* This person has defined their act as cheating, thus activating their belief that "Cheating is bad." But for them, this is not an iron-clad belief—and somewhere above it in their hierarchy is the belief that "Bad things aren't as bad when everyone else is doing them." This belief essentially gives their brain permission to cheat under these specific circumstances, even though they would agree that they're cheating & that cheating is generally bad.

This reasoning might be replaced in other versions of this belief-based rationalism by beliefs like "If it doesn't hurt anybody else, it's not wrong" (which *grading on a curve* or *admiration for a teacher* might negate) or some version of the very simple & effective belief "I'm special—these rules don't apply to me." No matter how they rationalize it, this person is likely to feel at least some guilt over their act, but they can live with it.

Belief reliance: *I'm not using my notes. I don't care what everyone else is doing—that would be cheating.* This person is likely confident enough in their belief system that they are less prone to use rule-based rationalization in order to achieve a short-term gain. This confidence also likely makes them less prone to have an imprecise, but convenient belief-hierarchy in which over-generalized beliefs like "Bad things aren't as bad when everyone is doing them" end up as top-level beliefs (which is potentially very dangerous).

This person has been conditioned to feel that the best strategy is the application of strong, specific beliefs to brutally-accurate narrative construction. Guilt obviously isn't a factor here—and neither is the pain that can sometimes result from the perceived "loss" of an unexploited value gain like cheating. As desirable as it seems, this belief-confidence (which often results in socially-constructive behavior) can also get...*ugly*. If your belief system has, for example, over time been able to convince you of the absolute inferiority of certain other races & you've developed a naive overconfidence in these beliefs—well, in these kinds of cases strict *Belief-Reliance* clearly begins to show some of its potential flaws.

Belief confinement: *Cheating would totally help me, and everyone else is doing it,* *and it's not like it's gonna hurt anyone, but... what if I get caught? I just can't. I'd feel too guilty*. This person's brain (likely because of previous behavior-patterns) has used their narrative construction to give them several possible reasons to either violate their "Cheating is bad" belief or apply a higher level belief. Alas, their belief that "Cheating is bad" (and its prediction of possibly-dire consequences) is powerful, and it has confined their actions even in the presence of strong narrative motivations.

This is, of course, exactly what beliefs are *supposed* to do. Even though this person is likely to experience some of that "loss" pain from an unexploited gain, they're willing to suffer that pain instead of the guilt. And in more extreme versions of these *Belief Confinement-based* inner-conflicts, an individual's capacity to overcome that predicted & ongoing "loss" pain—in order to "make the right choice"—is heavily influenced by those endorphin-based *willpower* mechanisms discussed in essay #2, which are designed to aid us in exactly these kinds of *opposing-impulses* scenarios.

Belief Confinement can also be at the root of a student's choice to cheat in order *to not look like a square* ("Being popular is more important than anything" or "Being unpopular leads to misery"). The difference between this kind of narrative construction & *Belief-Based Rationalism* or *Belief Failure* is the goal of the behavior that the belief is "confining" or "rationalizing" or "failing to mitigate." In our rationalization & failure scenarios earlier, the student wants the gain of *a better test score;* the rationalization allows them to use a higher level belief to achieve the desired gain & the failure allows them to essentially *ignore* their beliefs. In the *to-not-look-like-a-square* confinement example, the student might actually prefer *not* to cheat (making the act of cheating feel more like a *loss* than a gain) and yet might still feel *compelled* to cheat (or *confined* to cheating behavior) in order to adhere to their powerful beliefs regarding what is socially acceptable in high school's uniquely-convoluted communal structure.

In other words—*Belief Failure*, *Belief-Based Rationalism & Rule-Based Rationalism* are all ways in which our brain chooses to violate a belief in order to pursue a gain (or avoid a loss). Oppositely, *Belief Reliance & Belief Confinement* are ways in which our brain chooses to adhere to a (usually strong) belief in order to refuse a gain (or accept a loss). Basically, in the properly "confining" hierarchical combination—beliefs can be used to *make us* do pretty much anything (just as *Belief & Rule Rationalism* can be used to *allow us* to do pretty much anything).

This does not mean, however, that in those confinement scenarios our beliefs are an essentially uselessly-relative & socially-manipulative tool. In truth, I think most of us make our most-difficult "correct" (most ultimately-beneficial) choices in this *Belief-Confinement* way—not in the swaggering, defiant fashion of the *Belief-Reliant* person. (And in human behavioral terms, flexibility is often the *most preferred* trait in a system or the state in which its “equilibrium” is most sustainable—*adaptability* being our primary evolutionary advantage.) Usually, when caught in the grip of a "tough call," we are wanting oh-so-badly that delicious in-our-reach gain, and are only kept from it by some annoying, nagging *behaviorally-confining* belief.

Which sometimes makes us wish that we *didn't have* those annoying beliefs hanging around and killing our buzz. But after this final example, you might feel differently. This one doesn't really *belong* in our examples (because it's based on a brain with an inborn deficit) but it *does* occur in some cases. And this outlier powerfully demonstrates the importance of beliefs. Plus, it's pretty fascinating—in a somewhat *disturbing* way...

Psychopathic behavior: *I'm obviously using my notes, because it'll help me & I probably won't get caught, and if I do get caught, I'll just point out that everyone was doing it, so she'll have to punish all of us, which is almost the same as punishing none of us.* This is what you get when you don't have a functioning belief system *at all*—which likely leads to the development of a more-robust rule system (in order to help create more-reliable complex predictions in the absence of prediction-aiding beliefs). Our theory hypothesizes that this non-functioning belief system is the primary neural deficit that is at the root of most psychopathic behavior.

A psychopath's lack of belief-invoked guilt or remorse, their tendency to be capable manipulators (a likely result of that over-compensating rule-development) and their focus on the pure value-propositions in every situation regardless of the situation's societal (belief-defined) "moral" constraints—these are all hallmarks of psychopathic behavior. And you can create all of those effects simply by shutting off someone's belief system.

Thus, it’s a mistake to call psychopaths emotionless (as they are often described). Even without beliefs, the rest of their emotions can still function. This means that they can use them to make calculations about value gain/loss, predictions, and Agents of Value—which are crucial to that effective manipulative streak. And they display (and appear to feel) plenty of emotions: anger & rage (often apparently uncontrollable) over a loss, animosity towards potential Agents of Loss, gratitude for a gain provided, selfishness surrounding their own resources, pleasure over some machiavellian success, excitement over anticipated gains.

And it makes sense that some of the more evolutionarily-weighted emotions (like anger) would be expressed most readily & perceivably—considering these individual’s lack of behaviorally-calibrating beliefs. Additionally, emotions (& brain areas) that are closely related to & often accompany disgust (like fear) might grow generally weaker in psychopaths—like a muscle that *under-develops* due to the total absence of those frequent disgust-related usages. (Keep in mind that *every time* we experience the disgust or guilt of belief violation, that judgement is predicting that the behavior is likely to lead to an eventual *loss*—which automatically triggers *fear.*) Regardless of how outwardly muted or powerful these emotions may appear in any particular psychopath at any particular moment, it's likely the emotions (and their necessary calculations) are in there somewhere.

Therefore, they're not always "faking" these emotional displays (although they often likely are). But—although psychopaths can *judge & feel* these emotions within themselves—much of our *behavior toward others* is guided by our belief systems. (Look at how people from different cultures or families are conditioned to *believe* that affection is expressed with different types of behavior—which is the source of much marital distress.) Thus, despite feeling the emotion, a psychopath may show little outward display of their feelings if they don't deem that behavior as helping them to get what they want in the moment.

They could calculate this decision using advanced rules, which—unlike beliefs—would only likely orient the behavior from the perspective of the individual's personal gain. In other words, a psychopath's human interaction is primarily a result of a pure self-value-based emotional calculation; the attendant behavior may or may not be necessary in their rule-based view.
And their success in manipulating others—in "playing" people to achieve their gains—directly contradicts another common misperception about psychopaths: that they lack empathy. Empathy is a function of our mirror neurons, and mirror neurons play a key role in our ability to perceive (& subsequently *manipulate*) the feelings of others. In addition, mirror neurons play a key role in lots of other *and much more fundamental* processes—like *language acquisition*. This means that if psychopaths were really suffering from "abnormalities" in those *mirror-neuron-based* empathy mechanisms, they'd display a lot of other much more apparent & developmentally-altered behavior than simply *behaving like assholes*.

Narrative Complexity actually hypothesizes that mirror-neuron-related dysfunction is at the root of two closely-linked [21] —but according to our theory, oppositely-caused —neural conditions that we'll discuss in more detail later: autism (overstimulated & indiscriminately-applied mirror neurons) and Asperger's (non- or low-functioning mirror neurons). Because mirror neurons (typically) are devoted to specifically identifying & analyzing other "like entity" data input, effectively perceiving how someone is reacting or feeling and then faking the appropriate response to achieve your gain *requires* empathy (in addition to strong rule-based prediction skills).

Unfortunately, even if you still have the ability to feel someone else's loss—and you're good at rationally understanding that their loss *isn't* actually your loss—but you don't have any beliefs that define pain-infliction as bad, then empathy can't make you a "better" person. (Even healthy, empathy- & belief-capable people who simply don't *believe* pain-infliction is always bad can make great & almost-guiltless torturers.)

In addition—in situations like cheating or stealing or murdering—beliefs are what help us to know when an obvious value gain or loss avoidance is better to be left alone (for some bigger reason than our own individual benefit). And beyond just teaching us *when & how* to express emotions like affection, beliefs are also what *compel* us to behave in those ways that express our affection (because we are normally driven to avoid the guilt of non-compliance).

We all learn how to best show our affection through whatever social group we are in, and we feel compelled to behave accordingly — judging how much affection we have for someone and calibrating what has been learned to be the *expected* response. Thus, we hug someone we view as a high potential Agent of Gain because that's what our beliefs tell us we *must* do if we've defined that person in this way (*if you don't hug your mom, you feel guilty*).

A psychopath can still judge someone as a potential Agent of Gain, but if there is no purely narrative reason to hug them at that moment (e.g., I want them to give me a cookie right now and hugging will help) then they aren't compelled to hug that person because they have no behavior-guiding belief that compels them to hug them *just because they have "affection" for them.*

Furthermore, even though they can technically have that affection for a person, they don't *feel* it in the same way that most of us do. That's because most of us accompany our pure potential-value-based affection with something else: *admiration*, which is an emotion that relies on beliefs. Consider this: *a son has a father who gives him everything he wants, but the son knows that his father murders innocent people to earn a living.* If this son is disgusted by his father's behavior and thus, does not *admire* him, the son's overall feeling of affection is likely not very high (or at least it's *conflicted*)—despite his dad being a high-value potential Agent of Gain.

It seems that without *admiration*, a child's love just doesn't have that same *shine.* Which is good description of how psychopathic children appear to feel about *their* parents. (In addition, since we learn so many of our beliefs from our parents' behavior, we are more likely to admire them —and acquire that *shine*—due to those common beliefs.)

This diversion into our darker brethren tells us one thing above all: beliefs are fundamental to a healthy human existence. But I think it also tells us something else: psychopaths are not inherently "evil" individuals. They haven't replaced healthy, productive, non-violent beliefs with some opposite, socially-destructive set of beliefs (which is a case for epidemiologically separating the now nearly-synonymous terms psychopath & *sociopath*—since the latter well-describes individuals whose systems all function, but whose beliefs are simply *totally screwed up*).

Psychopaths merely view the world as truly self-centered beings. All gains & losses are about *them*. And as they grow older, they essentially remain an emotional infant, but achieve the logical & perceptive capacities (and needs & desires) of an adult. Combine this with other aberrant behaviors that are likely to result from an out-of-control rule system (grown hulk-like in its lifetime of overcompensation & overuse) and you have the blueprint for *dangerous* psychopaths like serial killers.

Those aforementioned out-of-control-rule-system-based aberrant behaviors can include troubling stuff like: *ritualism*—ingrained & repeated rule-based behavioral "causal sequences" containing excessive, non-essential actions that are incorrectly perceived to be necessary in order to achieve the sequence's intended result; *fetishism*—ingrained & persistent need for specific pleasure-seeking acts (like sex) to be accompanied by highly-specific rule-based criteria in order for those acts to produce actual pleasure; and *extreme behavioral rigidity*—ingrained, persistent & inflexible adherence to one's personal rules regardless of the behavior's impact on others, and a rigid unwillingness to violate or compromise one's personal rules at the request of others, regardless of circumstance or social expectation.

So, *yes*, this is a combination that's very likely to very quickly produce very undesirable results, but that is not *necessarily* pre-destined based on the neural deficit. I believe that early intervention (toddler-age) with a focused program of rigorous, specifically-applied rule-building would help to make these people much more functional in society. Unfortunately, it would be awfully hard to be certain that undesirable results wouldn't eventually emerge. In the end, without our beliefs, human brains just don't work very well (that is, if a healthy social fabric is one of your goals).

Leaving behind our tangent into strangeness, and returning our discussion to *all* of the belief- & rule-application examples we've discussed—what do they *collectively* ultimately tell us? They tell us that when it comes to decision-making, our use of beliefs & rules to structure or frame that decision is highly flexible. They also tell us that *how* we structure that narrative and *how* we've prioritized our beliefs ultimately define every conscious (or *non-reflexive*) decision we make.

But a cognitive process like the one presented here also raises a question about those beliefs & syntax: if our belief resource is applied *after* syntactic narrative construction, how can we include conscious & verbalized consideration of those beliefs in that *pre-belief-resource* narrative-construction location in our loop? In the view of our theory, the answer to this is reflected in the way that we build our vocabulary resource from other occurrences of those words stored in our memory (& from the emergence of those words & their associated data in our "working" memory).

This kind of "dual-presence" in our right-brain memories & our left-brain cognitive resources is also true of beliefs. Those experiences in which we've been told a belief by others or thought about it ourselves is the data that's the *root source* of any belief that ends up in our cognitive belief resource. And some of those experiences (& thus the word-based narrative parcels that express those beliefs) are recorded long-term in our memories, making them available to be used in our narrative construction when situationally appropriate.

But merely *expressing* a belief in this way does not mean our behaviors or actions will automatically adhere to or be impacted by this *expressed* belief. That's because that behavior is determined by where this expressed belief *actually resides* in our belief resource hierarchy—aka, the belief's *strength*. This means that if our action or our syntactic definition of that action actually violates a belief that is *stronger* than the one we internally or verbally expressed, we might still hesitate to act or might feel guilty about the act even though it does not violate the *expressed* belief (which was expressed *instead* of the stronger belief merely because it was the first related emergent data to earn a slot in that particular round of narrative construction).

For example: *In the middle of a chaotic & un-policed protest march, your anarchist buddies urge you to throw the rock through the bank window, and you yell, "You bet I'm throwing this rock! The plutocracy must be attacked at every opportunity!" But in this same moment, as you cock your arm to throw the rock, you feel the urge to hold back, and suddenly your internal dialogue is filled with thoughts about what your mother would think. Next thing you know, you're dropping the rock.*

In our model, the rock-throwing hesitation could occur before the thoughts of your mother emerged (that momentary pause was what allowed you to generate them). Thus, the pause was actually the result of your intended action *violating* a very high level (but not yet consciously contemplated) & *bourgeois* belief like "Vandalism is wrong"—causing you to hesitate even though the action was strongly supported by your actual syntactic construction & your *expressed* belief (*and* your desire to look cool in front of your fellow anarchists).

Basically, we can *say* that we believe anything that we want or think we ought to believe, but beliefs are a *very real* thing—a specific & powerful element of our cognition. In other words, for our brains to actually guide our behavior according to a belief (e.g., produce guilt when it's violated) that belief *must have*—through experience or study—actually *earned* its place in our belief resource hierarchy. Thus, your capacity to identify & articulate a belief via memory-based data is not the same thing as actually having that belief filed & applicable within our belief resource.

This capacity to identify & articulate a belief via memory-based data *does*, however, explain how can we include conscious & verbalized consideration of those beliefs in that *pre-belief-resource* narrative-construction location in our loop. Which means, having answered our aforementioned question, we can move on to our *next* declaration about these narrative-building architects within us...

All Architects Are Not Equal

Here's something that's pretty obvious about humans and their rules: some of us display a greater capacity for handling, building & applying these rules. Generally speaking, this capacity appears to be pretty-well hardwired in us from birth. We'll take a broader look at this kind of *nature vs. nurture* in our brains near the end of the essay, but since we're going to talk about the hardwired capacity of our rule systems—essentially, our *intelligence*—we'll catch our first glimpse of *nature vs. nurture* here.

Current theory generally divides intelligence into two categories: "fluid" & "crystallized" (*their* terms, not mine). Fluid intelligence—long believed to be a fixed, life-spanning attribute, aka *nature*—is equated with "pure" reasoning, logical thinking, problem solving, pattern identification, etc. This is what IQ tests are intended to reflect.

Crystallized intelligence is considered to be a capacity to apply learned skills or information. Although most theory does not generally not *equate* crystallized intelligence with memory, it is, nonetheless, supposedly *reflected* by one's accumulated "general knowledge" or vocabulary. (Just exactly *how* are they able to explain why something would be *reflected* by accumulated knowledge, yet not actually *equate* to that knowledge? What adherents of this theory are intuiting is that the way in which we associate & organize our *rules* affects how we apply that *accumulated knowledge*.) Unlike fluid intelligence, crystallized intelligence is *not* considered to have a fixed capacity—aka *nurture*.

But a new chink has been found in the armor of fluid intelligence's supposedly *fixed* nature: recent experiments seem to have proven the ability to improve fluid intelligence through the practice of very specific mental tasks.[22] This practice (which must be done intensely & regularly to yield any results) typically involves something called *n-back* tests, which essentially provide practice in quickly remembering & matching items from a previous set of items in a sequence (the tests grow in difficulty as they progress). The subsequent increases in IQ scores are not huge (this isn't *Flowers for Algernon*) but *any* improvement in fluid intelligence appears revelatory in the eyes of most current theory.

However, in the view of Narrative Complexity the results of *n-back* practice are not surprising. Just as the flaws of a short-term memory cache are easily ironed-out by applying our preferred looping mechanisms, I believe those same mechanisms handle "intelligence" with greater elegance than the currently dominant "fluid" & "crystallized" models.

So, in the view of Narrative Complexity, is there a *fixed* inborn aspect of intelligence? Yes. In fact, there are several. But these fixed aspects aren't limited to the area of cognition ("fluid" intelligence). Likewise, the trainable aspects of intelligence are not limited to our areas of recall & association ("crystallized" intelligence). Yes, the effects of our inborn capacities have a much different *impact* on each of these systems, but this is mainly a result of each system's specific mechanics (its *use* of those inborn capacities) not because the capacities of one system or the other are *wholly fixed* or *wholly trainable.*

According to our hypothesis, the inborn elements that most impact all of these systems are likely the same: our individual neural networks' data & associative capacities, the strength of those imprinting systems, and the speed at which it can process data. But, as we said, the effects of these inborn capacities are very different in our narrative-building mechanisms ("fluid") and our data storage systems ("crystallized"). In our data storage, greater inborn capacities can result in things like a better memory (longer & more storage, more reliable recall) and a greater ability to usefully associate unlike ideas (likely achieved both through better processing speed & greater associative capacities—major factors in creative insight). Nonetheless, all of these abilities can be strongly improved through a couple of simple methods: study & practice.

Even if you have a greater *ability* to remember lots of data, you can't make much use of that ability if you don't actually *feed* lots of data into your brain. Conversely, even of you have inborn limitations in data storage, you can still store & access huge volumes of useful data by feeding lots of it into your brain and using learned memory techniques (like narrative) to help you remember & recall that data. This makes the usefulness of our data storage systems highly-malleable even despite our fixed inborn capacities.

In our data storage, the main mechanism that our brain uses to overcome those inborn limitations (in addition to applying memory devices) is that essential memory mechanic:

repeated recall. Repeated recall can help to make-up for those deficits of a weak imprinting system & slower processing because it helps increase imprint strength and the fluidity between associated data. These mechanics (and those leading to a more-organized rule-set) account for the "improvable" mental capacities associated with that (*hopefully-being-debunked*) "crystallized" intelligence.

Improvement of our narrative-building mechanisms, however, is more restricted by the fixed inborn capacities of our neural network. The main reason: that repeated recall is not very useful in improving those fundamental narrative-building mechanisms. IQ tests, therefore, tend to reflect those more *fixed* neural capacities because they essentially judge the kind of fundamental rule-recognition/application process that repeated recall does not enhance.

Why isn't repeated recall very useful here in making-up for our inborn limitations? For starters, this is one of those brief moments in the loop where our imprinting capacities (which can be enhanced by repeated recall) likely have little impact on the mechanism. Just *before* we build our narrative (back in that data storage maze) imprinting capacity is obviously important. And just *after* we build our narratives, each narrative's emotional output partly helps *determine* that imprinting capacity.

But *during* the actual narrative-building, imprinting capacity plays mainly one role: it helps us determine rule priority & make some rules stronger than others (within that *learned-rule* resource). Thus, someone with a greater inborn imprinting capacities might begin to apply a learned rule after fewer rule-building experiences than a weaker imprinter. Nonetheless, a weaker imprinter can still effectively learn & prioritize that rule via those imprinting-enhancing repeated recall mechanisms like study & practice.

Unfortunately—as mentioned earlier—this doesn't help in something like an IQ test, because that test isn't actually asking our system of learned rules to discern & build patterns. Rather, it's asking us to recognize & apply unique patterns that are demonstrated within the question itself—tasks that rely heavily on those *inborn* fundamental pattern rules. This kind of genetically-defined skill-source is also the reason behind some people's innately-greater musicality: because our basic musical rules are an individually-inborn resource.

Although study & practice can still help us to *learn* & *internalize* new rules over time (and help turn an innately mediocre musician into a better one) once a rule has been learned & internalized, the benefits of practice likely have little impact on how efficiently we ultimately *apply* all those rules (which is why, no matter how much you practice, you're never going to play music like *Prince*). That's because the ultimate efficiency of rule-application is generally governed by our inborn pattern & data processing abilities.

And when no *learned* rules are used, rule application is governed by that innate ability to efficiently recognize, compare, analyze and apply patterns in the construction of a *unique* (aka, *unlearned*) response—i.e., to provide an answer to pattern-problems like those on IQ tests (which judge something different than the memory-recall & association processes judged by a test of factual knowledge & learned rules).

Its heavy reliance on those inborn capacities & rules (and the absence of repeated-recall's benefits) make this fundamental rule-recognition/application ability awfully difficult to improve. But those recent *n-back* experiments have shown us that there's at least one way to improve this ability (although the effects are short-term & it's unclear whether or not those limitations can be overcome).

How do *n-back* tests help to achieve this IQ improvement? I believe these *n-back* tests teach us new rules that help us to apply versions of those "data maximization" techniques to rule-application. These new rules are so fundamental (but unique) that they can be broadly applied to the actual *process* of rule-application. These would likely be rules about how we arrange patterns most efficiently in order to increase data resolution & therefore conduct more complex pattern comparisons using the same physically-limited systems.

And the reason that *n-back* tests improve *most* people's performance is because these are such unique & typically-unnecessary rules that few of us ever find a way or need to learn them. Thus, the benefits appear across almost all demographic categories. In addition, the way in which these *n-back* tests are administered is what helps even individuals with lower capacity neural systems learn & apply these new rules: lots of intense practice. Here repeated recall makes its single contribution to rules: helping to imprint new rules & make them stronger. Once we've learned (via intense *n-back* training) this new *rule-maximization* rule, we can use it to slightly enhance our limited inborn rule-application capacities.

And the *temporariness* of the IQ improvements in these experiments is fairly predictable in the eyes of our theory. *N-back* tests aren't likely impacting our inborn, baseline rule-recognition/application ability—they're just providing us with a super-efficient rule-maximization rule. The problem with this unique new rule: in everyday life it's not very commonly *useful* (thus our unfamiliarity with it).

Once someone has stopped regular *n-back* practice, they don't actually apply these new rules in their lives. Therefore, they're no longer benefitting from the repeated recall that helped our *n-back* boot camp make these new rules so powerful & frequently-applied. Now when they take the same IQ test, those much stronger, less-efficient, but much more commonly-used inborn rules are applied *sans-maximization* to the pattern problems. Viola! We just got *dumber*.

But did we really? The fact that we soon stopped applying those rules tells us one thing about them: they're not very useful *in our actual lives* (which is why almost none of us ever learned them in the first place). Therefore, the *useful application* of our "fluid" intelligence—which is all that really matters—is not exactly the same as what an IQ test might be able to gauge. Although *n-back* training improved IQ scores, the impracticality of the new rules made them essentially *useless* in everyday rule-application—basically making the IQ improvement a reflection of nothing that truly matters. In fact, we could spend an entire essay talking about the *true* definition of intelligence. But we're smack-dab in the middle of another essay already, and we should probably get back to it…

Our Inner Theater

These matters of intelligence—and the rule-building, recognition & application that helps define it—are all mostly about how our brain *uses* memory data, but there are still matters left to discuss about that memory data itself. Matters such as our actual *experience* of consciously recalling memories. The most fundamental enigma about the experience of remembering: *what exactly are we watching in our heads?*

Usually when we retell an old memory to others or ourselves, we experience the sensation of seeing this memory play out in our minds—like a little inner theater projecting short films from your past. (Unless you suffer from the *inner imagelessness* of the disorder *aphantasia.*[23]) How does our brain manifest such a depiction? Our visual systems are immensely complex (a result of that ever-increasing importance within our vertebrate lineage) and from our meekly human point of view, the results are nearly *magical*—although research assures us that there is, indeed, nothing magical about it.

From our theory's view of this process, part of that *near-magic* is its ability to "superimpose" very faint images produced from internal dialogue data essentially *on-top-of* (or along with) that much more visually dominant & pristine *actual* environmental data. Because our "Dynamic Core" actively integrates multiple data sources via our pre-frontal cortex in the production of our conscious experience, once our internal dialogue (& its attached *memory-based & very low-fidelity* sensory data) enters that arena, it has fleeting access to those visual systems required to conjure that faint flicker of a narratively-produced image.

The typically extreme *weakness* of this narratively-produced image is why it helps to close your eyes or stare blankly downward when trying to replay these little movies: doing so cuts down on the amount of competing incoming *actual* visual data (closing your eyes) or lessens the attention devoted to competing incoming *actual* visual data (staring blankly). This helps to gives that timid memory-based visual data a fighting chance in its ever-losing battle for our visual resources.

Even when you're retelling a story that you've heard from & happened to someone else, you likely have one of these weak visual depictions running in your mind as you tell the tale. Take a moment to do it yourself: first retell in your mind a quick (but preferably old and not that important) memory from your own past, then follow it by retelling a quick (but old, not important) tale that happened to someone else. I'll wait...

Okay, now think back to those two retellings and ask yourself: were the movies in your mind substantively different in quality? Did your own memory appear in HD while the other only had the quality of a VHS tape? Not likely. More likely is that they appeared roughly the same in your head. *But how could that be?* Isn't one based on actual visual & experiential data while the other is merely a re-constructed imagining? I have some more news that might disturb you: I think they're both essentially re-constructed imaginings.

Once upon a time, your own memory might have been of superior quality, but (assuming you retold an old memory, like you were *supposed to*) this far down the line, that higher resolution has long faded away—primarily a result of that ongoing memory degradation. As proven by our own memory's likeness to the replaying of the other person's story—just because we can "see" a memory in our heads does not mean our inner theater is depicting an actual visual recording of the data.

What happened to our high resolution data? And what are we seeing now when we replay those old memories? What the hell is going on, am I imagining everything? Actually, *sort of*. Look at it this way, those depiction mechanisms that use memory-based visual data to generate the images that we "imagine" are the same mechanisms that use ocularly-received visual data to generate the real world images that we "see." Thus, in both cases, what we're actually *perceiving* is a Dynamic-Core-generated *model (inside our brain)* that is based-upon & integrating all of those varied sources of visual (& all kinds of other sensory) data.

Our brain builds (imagines) our visual depictions based on the data available. In our consciousness viewfinder, the world we see is of extraordinary detail because the data input system (vision) and its gush of visual data is directly connected to our viewfinder depiction system. The data available is robust & the system has evolved to perfectly match the data input to its depiction. This is, after all, the depiction system's primary job, and these two systems have likely been working together since the time of lampreys.

In contrast, the memory storage system in humans and that visual depiction system are slightly odd bedfellows. Memory storage basically needs to use just a small amount of the depiction system's resources in order to help its data represent this key (visual) element of a moment. And our recollections don't really require those full HD viewfinder depictions. More to the point: they couldn't create them even if they wanted to, because those memory modules don't have nearly enough storage capacity to contain that full gush of visual data we consume in a "real" moment.

This is something we discussed in our essay about dreams. When our memory data is the source of visual depictions, the results aren't particularly impressive. And if we compare the two—visual elements in our dreams and in our old memories—they seem to have essentially the same qualities.

Some might mistakenly perceive this concept of a "consciousness viewfinder" depiction as flawed proof that there is within our minds some sort of "homunculus" (a silly-but-persistent philosophical notion that there is "someone" or some essentially metaphysical "self" in our mind that "views" these brain-painted depictions). What I'm intending to describe here is the rich & constantly "refreshing" visual data input that is integrated into a sustained dynamic multi-sensory neural model (again, Edelman's "Dynamic Core") whose multi-sensory data is subsumed & analyzed (in a priority-based fashion) by our cognitive systems, which allow us to consciously "perceive" & respond to data presented in that dynamic multi-sensory neural model. (If it sounds like I'm splitting hairs, it's because *I am*—but the mechanisms of consciousness are definitely a locale where hairs need to be split on occasion. We'll split these specific hairs even further in our *Hard Problem Addendum.*)

The purpose of such a dynamic neural model—and the reason why our wildly complex & fluid consciousness viewfinder ultimately emerged in vertebrates—goes back to those *lampreys* and their clever, new capacity to integrate multiple data sources (visual data & electro-sensory data) in the construction of a unified & fluid internal depiction of their nearby environment. By using multiple data sources to achieve the same goals (essentially, depicting & tracking objects) they were able to produce more detailed, accurate & data-rich 3D models of their environment. In order for these multiple & varied sensory data sources to achieve this kind of complex, dynamic depiction there must be *some* neural arena in which this simultaneously (& *rapidly*) arriving varied data can be integrated into a unified model—*aka*, some primitive, rudimentary version of that *Dynamic Core.*

This neural arena is necessary because the *ultimate* goal of this whole process is for the creature to actually physically & appropriately *respond* to what's depicted in their environment. And in order to respond effectively (which, in part, involves *predicting* where something might move *next*) that simultaneous, varied data must be *sequentially* processed in both *temporal* & *spatial* terms. Thus, a dynamic neural field aids this process by helping to unify simultaneous, varied data sources, and then by using those unified neural "moments" to create sequential depictions that track (& in later creatures, *record*) some of that data (*spike* data, which engages a creature's "attention") both temporally & spatially—which is necessary for accurate predictions & physical responses.

These are the roots of our own human *consciousness viewfinder*. And although our highly- & exquisitely-evolved Dynamic Core hardly resembles its early, rudimentary appearance in lampreys, *all* versions of this neural arena in vertebrates serve those same core purposes described above. However, it's important to note that just because data *appears* (& is integrated) within this dynamic core/consciousness viewfinder depiction does not necessarily mean that the creature will *respond to* (or *record*) that particular data. In humans, we might think of this as being *aware* of something without actually fully *perceiving* its presence (via our cognitive processes).

The data within this neural arena that creatures are most likely to respond to is that *spike* data, which garners more "attention" (*aka*, is more likely to be sent on to & subsumed by the next step in the data-analysis process). In humans—in addition to being driven by spike environmental data—this "attention" can also be powerfully, rapidly & continually directed, redirected & focused via our internal dialogue mechanisms. And the whole process of perpetually & rapidly redirecting that attention, and equally perpetually & rapidly employing that internal dialogue to note & respond to nearly anything & everything in our purview creates the wonderfully fluid *illusion* that we are actually "perceiving" *everything* that our Dynamic Core is technically *aware of*.

(The neural relationship between our awareness & attention, and how it shapes conscious experience, is well defined by Princeton neuroscientist Michael Graziano's pioneering *Attention Schema Theory* of consciousness—although the mechanism that he dubs "awareness" is actually what we label "attention," and vice versa. [24] Narrative Complexity further discusses how awareness/attention impacts data-processing in the next essay.)

Returning to our consciousness viewfinder's depiction of those *old memories*… If *your own* old memory looked the same in your mind as your memory of *someone else's* story (as it likely did)—what, then, are these images we see in our old memories? And where do they come from? Before we answer those questions, let's look more closely at those recent memories that seem to be in higher resolution. When we replay something that *just* happened, it still has that uncanny dream-quality in our heads, but it usually seems to contain more overall detail than a replay of an old memory (although it's still not an HD viewfinder depiction). *How is our brain doing this?*

I believe our most-recent memories have, essentially, higher resolution "media attachments" that can be temporarily associated with the word-based memory data. The reason why it's useful (therefore evolutionarily plausible) to have these temporary, recent high res media attachments is essentially the same as the reason why mundane dialogue hangs around in our head for a brief time before disappearing. Both mechanics help to give us that small window to "go back and get something" or give another pattern sweep to events that we brushed-off when they first occurred, but immediately require a quick recheck.

As we discussed when exploring pre-human mammalian cognition earlier (those "proto-narrative" structures)—basically, throughout evolution it's been beneficial for our brain to be able to provide a comprehensive answer to the question: *wait, what just happened?* This is likely because we often don't know the *real* importance of what just happened until we see the result—until *after* it happens. And if "*what just happened?*" doesn't arise quickly, our brain takes that as permission to continue the standard processing of our recent memory-data according to its initial imprint—which ultimately allows most of those recent (and low priority) high res attachments to fade away, leaving more generic attachments to do their job.

What exactly are these high res media attachments & this generic stuff ? The difference between these two goes back to associations & data resolution. When we're replaying one of those very recent scenes, its few specific narrative parcels don't have enough capacity in their modules to recreate in detail *every* visual (or other sensory) aspect of that replay. But its recentness means that there are plenty of easy-to-access (temporally-surrounding & closely-associated) memories that haven't faded away yet. And those memories might've focused on those other visual elements not contained in detail in the target scene. These other (likely only temporarily-stored) memory modules serve as high res media attachments: associated neurons that possess some of that relevant more-detailed sensory information.

Thus, when you replay those few, specific, very-recent narrative parcels (the scene) your brain can enhance the depiction with detail from that other closely-related sensory information—which is not actually temporally-simultaneous (and not actually contained in the target scene's few specific narrative parcels). Although none of our memory's version of high res visual data is nearly as robust as the HD viewfinder stuff our eyes process, using several of these focused-but-fuzzy object-depictions can help us to build a broader & more complete (higher res) overall scene than we can using the few focused-but-fuzzies that are contained in the target memory's limited narrative parcels. I know, *huh?* Don't worry, this example should clear things up:

Very soon after my wife came home, I replayed in my mind a specific moment of her arrival in which she walked up the steps & waved to our little girls, who were standing at the big front window. I could see the whole scene: the car she'd just parked in the street behind her, her expression & what she was wearing, what the girls were wearing as they stood in the window. Was it raining? Let me think...yes, it was raining lightly.

This very-recent memory seems full of detail. That detail, however, is likely a result of some *slight of mind*. We have been fooled into thinking we recorded all of these details in the actual scene's few narrative parcels. But these media attachments have likely been built from other surrounding moments that contained the richer detail of each specific element: *the moment when I saw my wife park her car, the moment I saw the girls run to the window.*

In the actual recalled scene—because my wife was the focus of my attention—the informational details of her expression & clothes might truly be contained within (or attached to) those narrative parcels' memory modules. And although the girls clothes & the car were likely ignored (or very *low res*) in the actual moment, during the surrounding moments—when those other elements were my focus—my brain recorded those images in more detail.

And when I think about the rain, *who knows where that data came from*—maybe looking out the window 10 minutes before. Nonetheless, adding it to the replay is a simple matter of the data being requested (by ourselves or others) and our brain judging that it has reliable-enough information to make the reasonable *assumption*, and quickly adding it to the replay. Even though they've come from other sources, these attachments' recentness (thus, their *undegraded-ness*) makes it all slightly more detailed and more convincing than that dreamy, old, unimportant memory I asked you to replay earlier.

If I try to recall the same moment several days later, it's likely that the "dreamy" quality has overtaken that high res memory. *What's happened now?* Now there's no longer any recent, related high res visual data—those media attachments weren't contained in important or retold narrative parcels and have since faded away. Now the replay must rely entirely on the scene's own few narratively-based parcels for its visual data. Because although none of the temporally-surrounding data has survived its half-life, I recalled this one specific scene several times—thus, it's still hanging around & accessible.

Now when I replay the moment, all I mostly have is that already-slightly-degraded specific image of my wife and the narrative framework: the words. And although those words aren't nearly as good as the real sensory data, they can still do the trick. That's because these words can help me to generate "generic" visual data when I replay the moment. So when the words "her car" appear in my retelling, my brain finds the most recent, reliable (thus most easily & likely-to-be-pinged) visual data for "her car" and uses that data to draw-up its dreamy version of her car in the memory retelling.

And if I continue to frequently recall this memory in this specific way, that particular generic car data might become essentially permanently attached to the original memory—which can be "re-written" little-by-little with each retelling, as the power of each new retelling slightly alters the memory imprints & structures, and their associations (or maybe even lays down an entirely new version of the memory, which eventually gets "first-ping" when the memory is called upon in the future).

Thus, every time I replay the memory later on, her car now continues to be depicted in exactly the same (but still dreamy) way. The eventual consistency of this generic attachment makes it seem like it was a part of the original memory. But it's simply placeholder data that became closely associated to that memory. Keeping this new attachment around long-term is no big deal because it's low-res & by now well-imprinted—thus having none of the drawbacks of the original high res media attachments from those recent, closely-related, but temporally-doomed memories.

Do I have any studies to support this hypothesis that memories are word-based patterns connected to recent high res media or generic attachments, which are primarily a re-imagining of the moment? *Not really*. Although—as discussed at the beginning of the essay—the latest neuroscience certainly indicates that our brains are very *capable* of (& possess the neural mechanisms required for) managing a system like this one. And I do, of course, have some personal anecdotes (very common experiences) that help illustrate these mechanics...

There's a memory I have from first grade that I have retold with great frequency. In short, it's a memory about hurriedly putting on my snowsuit & trying to get to the soon-departing school bus in time. When I tell the story, I can see it in my head: *Mrs. K's room, me leaning against a desk as an exasperated fifth-grader (our bus guide) helps me zip my snowsuit up, imploring me to hurry*. In my mind, the snowsuit is usually beige & hooded.
Would I be surprised if it was actually a beige winter jacket with blue snow pants &

a blue hat? Nope—that's entirely possible. In truth, although I *believe* it happened in almost the exact way I describe it—I wouldn't bet my life on it. What I *do* know is that whenever a related topic comes up, these are the words I generally use to retell the story that comes to mind, and these are the images I usually see in my head when I tell it. (In fact, I tend to see—as we often do—this ancient memory from a 3rd-person POV. In other words, I *see* my young self in the memory—which is an obvious tip-off that this movie is being re-imagined.)

This kind of common experience supports the notion that all memories are primarily language-based. Consider that most media attachments' detail—high res or generic—is dependent on what usable, recent, cross-matching sensory data is available to supplement our word-based memory *at the time of that specific retelling*. Therefore, if I'd retold this story when I was in 3rd grade, I might still have a high-quality, relatively-recent memory of that specific snow suit, and thus the memory's linguistic components "first grade" & "snow suit" would combine to ping a more accurate, detailed version of the suit.

By now, that data is long gone. All that's left are the words "first grade," "snow suit" & "beige," which are more than enough for my brain to create the generic visual item that I've seen in this memory consistently for the last 20 years. Accurate or not, *I still remember* the snow suit, the desk, the classroom. And for most of us, those three words—*I still remember*—are good enough. We'll battle 'til the cows come home in defense of something we *still remember*.

Of course, since each of us has memories that are essentially equally unreliable, your vehement belief in those memories isn't any less justified than the next person's. So go ahead, *swear* you remember. Nobody's really in any position to claim their version is more valid. (Unless, possibly, if that original event was simply *way* more important to them when it first happened—leading that virgin narrative to be both super-strongly imprinted and frequently, accurately recalled).

More Ghosts In The Machine

Depending on how closely your *own* mind was paying attention to our essay's recent data and your ability to apply the most appropriate syntactic rules—you *may or may not* recall that I promised personal *anecdotes* (plural) to support our hypothesis that memories are word-based patterns attached to sensory & emotional data. So, here's our *plural*. This anecdote helps to show just how powerfully word-based our memories are.

More than a decade ago my grandmother, in her early 80s, was diagnosed with Alzheimer's. As anyone who's witnessed their progression knows, Alzheimer's & other forms of dementia are diseases of exponential sadness. I often felt that my grandfather—who cared for her in their home & then visited her daily in her nursing facility until his end at 95—was living with a ghost. She was someone who might mystically, inexplicably, briefly appear out of the ether, then quickly disappear again into the shadows of the other side, unreachable. And that apparition of her—in a ghost-like truth—was usually temporally displaced: not perceiving or understanding the actual moment she was in, but arriving from and to somewhere else, a place only she could see.

In the time just before she moved to a nursing home—that final period when she was still occasionally almost-present—I spent an emotional, melancholy afternoon with her. It was the holidays and much of the family was gathered at my uncle's home. The occasion was, in fact, the last time I felt like I actually spoke to *her*—although she is still alive & this experience was almost a decade ago. During that afternoon she sat beside my grandfather on the couch, mostly with a contented far-off gaze. But every so often, some part of the conversation would spark something in her and a vaguely-related anecdote (usually from the long-ago past) would come spilling out in great detail.

There was one particular story that got caught in a kind of loop that afternoon. It was a story from her youth about a giddy weekend at a lakeside cabin with some relatives, and it included an aunt of hers who was a larger-than-life figure. I'd heard her tell it before in almost the exact same fashion & detail. And on that afternoon—after the story was brought to the surface in her mind—she told it not once, but multiple times, pausing briefly between each telling, then beginning again as if it had just come to her.

This is common behavior among Alzheimer's victims and other dementia sufferers. One of the extraordinary things about witnessing this kind of recollection is that someone who couldn't tell you whether or not they just ate that sandwich is suddenly able to fluidly, lucidly & *expressively* retell a decades-old story in great detail. Beyond that, in this case (as is common) my grandmother told the story several times in a row using *the exact same words*. Not almost the same words, but the *exact* ones. In addition, she inflected them almost identically, and accompanied them with the same facial expressions and asides—pausing at the same spots to provide the same details about the larger-than-life aunt, claiming each time how she could vividly picture the person or scene she was describing.

What is happening here? How is this brain's disrupted system able to recall such detail? And why is the detail so exactly identical in its depiction? As far as we can tell, Alzheimer's victims have developed a build-up of plaque in their neural structures. Basically, the plaque build-up inhibits our neural lightning storm. This not only cuts off lots of once-fluid data pathways among our right hemisphere's storage neurons, but ultimately disrupts our entire loop—turning a person into, essentially, a misfiring computer. Nonetheless, the brain is a persistent and adaptable machine. When one part is damaged permanently, it tries to reallocate resources and move necessary systems to a still-functioning area, re-shaping its self-built architecture wherever possible.

In situations like Alzheimer's, at some point this persistence is overwhelmed by the problem. But until then, that persistence can still occasionally propel a piece of incoming data to an actually relevant & still unobstructed-by-plaque memory. When this neural network is activated, these specific narrative parcels might (for any number of reasons) still contain good resolution. This allows the story to be told in an entirely natural & detailed fashion. If the brain can still find some path through the loop (which gives us access to speech & expression based on what's scripted into the data) then for a moment the ghost can come to life. In this moment of telling the person is suddenly there, back from their oblivion— although not still quite with *us*, but somewhere else. Nonetheless, that somewhere else is rich with detail and emotion.

And the linguistic & expressive exactness of my grandmother's (and other Alzheimer's & dementia victims') multiple retellings—their *verbatim-ness*—seems to support that primary hypothesis: memories are word-based patterns. My grandmother recalled the memory each time word-for-word because that's how the data was stored: *word-for-word*. Keep in mind that Alzheimer's sufferers at this stage seem to have great difficulty taking newly-processed emergent data and using rules to construct complex *now-related* narratives. This is why they are almost never *in the moment* with us even when the ghost seems to have brought the person back briefly. Their mind is in a semi-dream-state, primarily grounded in & generating "reality" from the memory data.

This is because their processing of the present is limited to the most basic *I am here now & you are with me now* depictions. This seems to be the most central & primitive state of consciousness—the loop running in some bare-minimum mode. This is likely akin to that original state of self-awareness around which most other complex self-building mechanisms evolved (just as our modern emotions evolved around those still-present proto-emotions).

Because only this minimum state is (occasionally) achievable in Alzheimer's victims (eventually disappearing completely) the actual complex temporal & circumstantial details of the *now* essentially cannot be narratively-integrated by the damaged brain anymore (except on rare occasions). Thus, the retelling of stories cannot be tweaked or embellished on-the-fly according the present moment's audience or circumstances.

These limitations help result in the unaltered exactness of the repeated retellings. Everything my grandmother said was likely pure, unembellished, from-the-old-neurons recall—and that recall seems to have been, at heart, all about remembering the words themselves. The specific words in a specific order, each accompanied by specific images and a specific delivery must have been included in (& were likely the foundation of) that old memory-based data.

And we can't explain-away this exactness with the recent-memory mechanics that might make just-spoken data newly high-priority & easier to recall exactly. In an Alzheimer's victim, that first retelling *couldn't* suddenly get seared as a complex, accurate recent memory—making it easier to subsequently repeat in lengthy detail word-for-word.

These individuals have mostly lost the ability to record *any* new memories. Their recording mechanisms may temporarily come online enough to record a few repeatable, looping recent parcels—*that's a lovely sweater*—but not likely enough to record & repeat a long, detailed, identically-expressive narrative. So the source of any exact repetition *must* be that old long-term data. And if it's repeatable as a word-for-word, smile-for-smile narrative, then the memory data must be—at its core—stored as a *word-for-word*, *smile-for-smile* narrative. *I know*, that's probably the third or so time I've tried to convince you, but I also know you have your doubts (reasonably) about memories essentially being the words we use to retell them. However, I think this final example is the most convincing—basically because there is no other decent explanation for my grandmother's (and other Alzheimer's/dementia victims') *verbatim-ness.*

Keep in mind that we did not design our memory solution *around* this Alzheimer's experience—our theory's memory systems are based on the needs, limits, capabilities, behavior & evolution of the brain. The fact that Alzheimer's victims' *verbatim-ness* is well-explained by the system we've already discerned hopefully just helps to *validate* Narrative Complexity*'s validity*. To me, it looks like further proof that the deeper you dig, and the more you connect our sub-systems' wide array of intra-cranial dots, the more sense our brain's entire elegant machinery seems to make.

When Good Brains Go Bad (or When They Get Unique)

As our exploration of the effects of Alzheimer's has shown us, it is often the saddest or most-troubling brain events that provide some of the most-unique windows into the mechanics of our minds. (*Ergo, the brilliance of Oliver Sacks.* [25]) And brain events that negatively impact memory & its surrounding cognitive mechanisms can create some of the most severe of these *deficits of mind*. Yet, troubling as they are, these deficits can help illuminate much about how memory & cognition work. In the view of Narrative Complexity, these neural deficits typically result from one of three general causes: system atrophy-cum-failure (aging), system disruption (injury, disease or inborn deficit), and system dysfunction (emotional or chemical imbalances). We'll examine these three deficit causal categories one-by-one.

First, *system atrophy-cum-failure as a result of aging*. It becomes a more obvious (and depressing) truth with each of our decades of existence: as we get older, those once awesomely-calibrated, highly-flexible, dynamic, resilient, easily-trainable systems in our body—from head to toe—start to atrophy or break down. Not only do we typically do big things with less grace—like run slower, jump lower, forget more often and see with less acuity—but humiliatingly little things begin to diminish everywhere (*we even urinate with less vigor, for goodness sakes—they forget to tell you that*).

Thankfully, it has been shown that sustained, robust & well-paced use of our bodies as we age can help to significantly slow this atrophy in many physical systems. And the biggest benefits of continued robust use seem to appear in the latter half of our lives, where such usage can essentially flatten the atrophy curve as we enter mid-life and can make our systems much more functional in old age. However, as demonstrated by even the most finely-tuned & hard-working elite athletes, that youth-to-midlife downward curve in functioning is essentially inevitable. Hard as we might try to avoid it, we are doomed to begin slipping from that maximum efficiency to that level where we can maintain a nearly flat-line decline after mid-life. And just as that youthful maximum efficiency is unsustainable, ultimately that mid-life flattened curve begins to degrade. As we push deeper into old age, that drop-off is likely to eventually become more precipitous.

Although science (& experience) indicates that our neural systems are among the most sustainable deep into old age—and the most positively-responsive to that robust use over time—like everything else in the body, its systems still ultimately atrophy, perform with less vigor, and fail.

If this decline mirrors our other systems, then our drop-offs in mental performance are primarily due to the physical mechanisms of our neural system weakening. The most apparent physical problem resulting in neural decline seems to be in the mechanisms producing our emotional juice & neural-loop energy. *(Egads, even our brain excretes with less vigor as we age!)* This likely contributes to the fact that—although we can still experience intense emotions in old age—generally speaking, even the most emotionally volatile of us tend to (as the saying goes) *mellow with age.*

In addition, the memory potential produced by these imprinting mechanisms is therefore less potent, leading to less efficient new-memory imprinting & recall performance. And it makes sense that the most-recent, likely mundane memories (those illusory short-term ones, which are already the weakest & quickest to vanish) suffer the most from this neural decline. This is why, as we get older, we forget stuff we were just thinking or just about to do *all the time*.

Such recent-memory deficits even seem to happen in individuals (like 80-year-old poets & professors) whose cognitive-processes remain extremely robust. Thus, it appears that this memory problem (fortunately) does not have a highly-detrimental impact on the use of already well-stored data and narrative-building rules. Therefore, we can still comprehend & tackle big problems with this typical neural deficit of age. (Although we should probably have a pen & paper handy as we calculate our solutions—so we can make use of that most primitive & rudimentary of memory-limit-circumvention techniques: *writing stuff down*.)

In fact, there are likely two opposing forces of aging that can make our minds both *less* and *more* capable as we grow older. Those atrophying imprinting systems & weaker neural connections (which lead to worse remembering & recall, and less fluid associative pathways) are obviously a detrimental aspect of aging. But older brains that have been well-fed & nurtured can also possess a distinct advantage: that life-long accumulation of deeply-interconnected data, rules, vocabulary & beliefs—which can lead to that calm, confident & assertive decision-making/problem-solving that exudes *I've seen all this before*. In other words, an aging brain can also mean a *wise* brain. (Some of these opposing forces of aging in the brain were recently explored by Mara Mather at the USC Davis School of Gerontology in her 2012 paper "The emotion paradox in the aging brain."[26])

Our second category of causes for these deficits of mind is not typically a wisdom-enhancer: *system disruption due to injury, disease or inborn deficit*. There are lots and lots and lots of ways for this to happen, and lots and lots of possible results. *(This is why you should, among other things, always wear a helmet and avoid inbreeding.)* But to illustrate this causal category, we'll first focus on some simple, common effects of general physical trauma (injury) to specific brain hemispheres. Two of the demonstrated results (in some cases) of these kinds of injuries: individuals with right brain trauma tend to make errors of *commission*, while individuals with left brain trauma tend to make errors of *omission*.

An error of *commission* is when someone gives (and believes) a nonsense description or explanation of a situation whose narrative elements are obvious to a healthy brain. *Someone with a right brain injury is shown a picture with a banana, a bowl of cereal and a carton milk, then told to explain the picture.* Essentially, they've been asked to create a narrative from these elements. Here a healthy brain is likely to say something like *pour the milk into the cereal, then spread some banana slices on top*.

But an injured right brain might, for example, have a hard time identifying the banana. This is because our data storage system typically appears to reside in our right brain. Thus, the banana-identifying memory data is garbled or inaccessible. As we said, there are lots of ways to mess up this system—but in one fashion or another, we've lost our ability to connect the incoming (environmental) banana data with the right-brain memory data used to *identify* the object as a banana. This makes it impossible to call-up the word "banana" (and its definition & use) from our still-working left-brain vocabulary resource when building our narrative here.

This injured brain works like a person with a limp, the hobbled right side forces the left side to do more work to *maintain reality* (narrative cohesiveness or validity). And the left brain contains most of those narrative-building mechanisms. So to make up for the unidentifiable object, it ends up "over-applying" some of its narrative-building rules. The left brain might view the banana more abstractly: *it's long and curved, with a narrow protrusion at one end*. Here the object has been defined by rules of physicality (patterns of appearance) not a vocabulary-based meaning. So it might try to derive the abstract object's use according to that rule-based definition, creating a seemingly-logical (to them) narrative. Thus, the right-brain impaired person might give an explanation like, *"Use that curved thing—I don't know what you call it—to open the carton of milk, then pour it on the cereal."*

This is an error of commission—they've made up something obviously absurd to explain & use the memory data they can't properly access. Forced to lean on their left brain's still-functioning narrative-building machine, they've created a dodgy narrative on-the-fly based on the insufficient incoming data. And yet, although it might feel a little fishy to the individual, they still believe it's a valid answer. This is because, not knowing what the object really is, there is nothing to indicate to them the absurdity of their narrative.

In fact, these right-brain impaired individuals might seem proud of their answers. They are prone to feel as if the pictogram & its one "unidentifiable" item is a kind of puzzle, and might take pride in the fact that they found any seemingly-functional use *at all* for the oddball item.

In contrast, errors of *omission*—which typically are associated with left brain trauma—are when someone can properly identify all of the pictogram elements, yet excludes the obvious use of one in constructing an explanation or narrative.

Therefore, if someone with this type of injury were given that same set of three images (carton of milk, banana, cereal) they might respond, *"Spread the banana slices on the cereal, then—I don't know—drink the milk while you eat it, but I guess you'd want a glass for it."* Here the right brain has done its job; it pinged all the correct relevant data about the objects and provided all of the syntactic elements necessary to create an obvious, likely narrative. It is not stumped by the banana, the cereal, *or* the milk—it *recognizes* all of them.

But the injured left-brain is misfiring, and its narrative-building rules & mechanics are no longer being applied efficiently or properly. Essentially, the machine has abandoned a piece of usable data (the milk), failed to apply an obvious rule (milk is added to cereal), and left the potentially-useful narrative-building element out of the primary narrative syntax. Here the healthy right brain likely isn't of much help (its job is mostly done by the time it turns over the data to the injured left-brain) but our mind might still try to solve this problem by tacking on some alternate, essentially narratively-separate use for the abandoned data—a story not very well interwoven with the primary narrative.

In the vast spectrum of brain injuries, disease & inborn deficit, of course, the ultimate effects of any system disruption do not always map so neatly to brain hemisphere & function. These systems are intricately intertwined, thus what appears to be faulty narrative-building might actually be something else in the system misfiring, causing an unforeseen cascade of effects that ultimately *presents* as a narrative-building error. This is why these types of brain-system damage & disruptions can produce such frustrating & mysterious problems. When you can't truly get under the hood to take a close look, it's easy to misdiagnose the real source of trouble.

Nonetheless, there is at least a general pattern to the results of certain types of disruptions—like the trauma-induced data-handling errors described above. And what we've at least shown here is that the evidence in these scenarios strongly supports Narrative Complexity's construction of looping mechanisms and the way that these mechanisms map to specific brain hemispheres.

Going beyond injury, we just discussed in detail a system disruption due to disease (Alzheimer's). And in terms of system disruptions due to inborn deficits, we already gave a whopper of an example of that too: *psychopathic behavior*. (Probably best not to revisit our dark brethren again at the moment—you never *really* know what might happen around those folks.) Since we've mentioned psychopathy, however, it seems appropriate to identify the neural disorder that (according to our theory) is essentially the *opposite* of psychopathic behavior: obsessive compulsive disorder (OCD)—the result of an *overly*-powerful and *indiscriminately-applied* belief system.

True behaviorally-disruptive OCD is marked by persistent, recurring & impossible-to-ignore compulsions & obsessions that carry a personal significance strongly (& illogically) disproportionate to the behavior's *actual* benefit. (*Read*: true OCD is *not* just a powerful-but-ultimately-frivolous *over-inclination* towards desiring neatness & order.) And in a cognitive system such as the one we've proposed here, all of those aforementioned behavioral symptoms could be produced by an overly-powerful and indiscriminately-applied belief system.

Consider that compulsions like unnecessarily repeating particular acts an exact number of times and obsessions such as washing one's hands after *every* possible exposure to infection are driven by the *belief* that *not* behaving in these ways is highly likely to ultimately lead to a bad result. In other words, OCD sufferers are constantly compelled to behave in illogical ways in order to adhere to their overly-powerful & indiscriminately-applied (i.e, inconsequential & predictively-ineffectual) beliefs. This dysfunction results in a kind of magical thinking whose dictator-ish control over behavior is, unfortunately, unmitigated by the magical thinking's illogic & inaccuracy. What, then, is the difference between these kinds of belief-driven compulsions & that rule-based *ritualism* mentioned earlier? A ritualist cognitively-but-unconsciously (& wrongly) *presumes* that they *must* perform this specific series of actions in order to achieve the intended (& usually specifically-defined) result of their ritualistic causal sequence. In contrast, a compulsive individual powerfully, consciously (& wrongly) *believes* that they *should* perform this (or these) action(s) because to *not* perform the action is—according to their belief—highly-likely to lead to an *ultimately* (& often broadly-defined or malleably-definable) bad result.

And when we examine the categories of compulsive behaviors that are commonly displayed by these *hyper-believing* individuals, it's not surprising that—in light of our belief system's strong connection to primal *disgust*, which is founded upon disease avoidance—OCD is often expressed in overly-powerful & indiscriminately-applied compulsions to engage in disease-avoiding behavior like hand-washing & other types of self-grooming or self-protective measures.

Although Narrative Complexity hypothesizes that the left hemisphere is the locus of the narrative building/analysis with which our belief systems interact—not much is known about what specific cortical areas are central to analyzing & employing those beliefs (our theory is the first to describe the particular kind of belief system proposed here). But since we know that beliefs are among the most-sophisticated & uniquely-human cognitive mechanisms, it isn't too much of a stretch to speculate that they make use of one our most-sophisticated & uniquely-human (& uniquely-great ape, -elephant & -cetacean) neural tools: spindle neurons.

Interestingly, one of the few brain areas where these spindle neurons have been located is the fronto-insular cortex (believed to be a key player in complex predictions & decisions). And it's been shown that the insula is highly-involved with processing that emotion that beliefs *rely* on: disgust. Thus, if we were going to venture our *best guess* at where to start looking for the roots of the neural dysfunction that results in psychopathic behavior & OCD, we'd venture somewhere in or around that *fronto-insular cortex*. (And research has shown that the brain phenomena that appear to correlate to psychopathy include diminished amygdala volume—which could result from the absence of those disgust-related *fear* responses aided by the amygdala—and dysfunction within that key neural disgust-processor: the insula. [27])

In the view of our theory (due to the insula's key role in managing disgust—whose modern emotional roots are closely tied to *embarrassment*) insula-related dysfunction is also likely a big player in another currently-mysterious neurally-based disorder—one that possesses a disturbingly high morbidity rate and a stubborn resistance to even the most intensive treatment: anorexia/bulimia. (Two disorders that are, according to our hypothesis, slightly different expressions of the same root neural dysfunction.)

Maybe the most vexing aspect of anorexia/bulimia is that it enables & encourages the one behavior that almost all chordates are inherently designed to avoid above all else: *starving to death*. Everything about chordate neural systems are, at some level, designed to achieve one ultimate goal: acquire & consume the resources necessary for survival (the most important resource being, of course, food). In other words: no matter what, eat something *or you will die*. What kind of dysfunction could subvert (continuously, often over the course of decades) this most primal & powerful of our desires?

I believe the answer to this question is hidden within that same vexing aspect: the ability to enable & encourage *not eating*—even when a hungry individual is presented with food that is clearly disease-free, ideally-prepared & deliciously-edible. There is, according to our theory, actually one unique (& brief) point in human evolution when human brains were likely *programmed* to avoid eating (apparently) disease-free, ideally-prepared & deliciously-edible food—*even when they were hungry*. This moment is the miraculous period during which our human ancestors (spurred by their control of fire) began to prefer cooked meat over raw meat, which (as discussed in Essay #2) led to the development of our visually-based *disgust* response toward raw meat and paved the neural roadway to our belief systems.

In our exploration of that evolutionary moment, we hypothesized that the primary behavioral mechanism & emotion that our ancestors used to socially reinforce that new, beneficial (but hard-to-achieve) *don't-eat-that-raw-yummy-wait-for-the-cooked-one* behavior was Pride/Embarrassment (*aka*, Inclusion/Ostracization). In other words, those human ancestors *shamed* each other into *not eating* (apparently) disease-free, ideally-prepared (to them) & deliciously-edible food—*even when they were hungry*.

This food-&-shame-based, socio-emotional behavioral mechanic likely served as a kind of evolutionary bridge between the early hominin brains that could not resist the desire to eat raw meat and those later, nearly-modern human brains that possessed (as we do) an inborn (visually-based) *repulsion* toward particularly bloody or "gory" raw meat. And it is not hard to imagine that before this *shaming-triggered-food-avoidance* completed its transformation into *visual-disgust-triggered-food-avoidance* there was specific neural circuitry designed to make these evolving humans more prone to *not eat* that desirable food in response to shaming. Eventually, those nearly-modern brains began to replace that *shaming-triggered-food-avoidance* with that more efficiently-applicable & reliable *visual-disgust-triggered-food-avoidance*—which could specifically encourage the *not eating* behavior in *all* raw-meat encounters.

And thus, in most modern human brains, that primitive, evolutionarily-short-term, *shaming-triggered-food-avoidance* circuitry is (if it still exists at all) a long-neglected, systematically-atrophied version of its once-powerful self. Essentially, if this vestigial ghost circuitry still even exists in most "normal" human brains, it no longer has enough privileged access to neural resources to have much impact on average behavior. Anorexia/bulimia, however, is the *opposite* of average behavior. And that's likely because, according to our theory, individuals who eventually develop anorexia/bulimia appear to possess a still anciently-powerful (or easily-revived) version of that primitive *shaming-triggered-food-avoidance* circuitry.

One of the things that seems abundantly clear when listening to the medical histories & personal stories of anorexics/bulimics is that nearly *all* of them can recount some powerful, preadolescent *shaming-based* experience related to their food consumption. In other words, at some point early in life nearly all sufferers of anorexia/bulimia were told by someone important (i.e., a close relative or a doctor) that they should *eat less food* because they were already or were about to become *fat*. (Or, in a smaller category of cases, individuals might've had some *other* powerful shame-based experience—like molestation—that essentially *cascades into* body-image-related & shame-based food-avoidance.)

Of course, not *all* people who have been shamed for eating too much develop eating disorders. In fact, although many, many people today have an "unhealthy relationship" with their food consumption habits (often due to socially-reinforced shame about their bodies) the vast majority still do not display anything close to a *life-threatening* capacity to refuse eating. Yet, as described, most anecdotal evidence suggests that nearly all true anorexics/bulimics have some powerful, preadolescent *shaming-based* experience related to eating. (And the subsequent behavior triggered by that shaming—behavior that seems to grow exponentially worse in adolescence—looks *exactly* like the kind of behavior triggered by our ancient proto-emotion *Ostracization* & its sibling-like descendant *Embarrassment*.)

This strongly suggests that it is not merely the food/weight-related social experience that is the source of this disorder. Instead, it suggests that a specific genetically-based dysfunction *pre-exists* in anorexics/bulimics and is *triggered* by the food/weight-related social experience. The dysfunction: a still anciently-powerful (or easily-revived) version of that primitive *shaming-triggered-food-avoidance* circuitry. And there is a unique problem posed by this old circuitry's reemergence in a modern human brain, one that wasn't present way back when it first came into existence: that old circuitry is now working *in conjunction* with those visual disgust & belief-based behavioral systems that long ago emerged from (and were *intended to replace*) those evolutionarily-short-term *shaming-triggered-food-avoidance* circuits.

This means that once this young, developing human has neurally (and, by dysfunctional accident, overpoweringly) defined *food avoidance* or *not eating* as top-of-the-list-high-priority behavior, that human's brain begins to calibrate its other behavioral systems in support of this newly vital (and survival-disadvantageous) goal of *not eating*. All of the sudden, those modern, ultra-powerful belief & rule systems come to the aid of this obsessive *not-eating* quest—setting up all kinds of behavioral fences that prevent the individual from doing anything that might interfere with achieving the unachievable thinness that has been defined (by that first food-shaming experience) as the socially-based reasoning for this all-important *not-eating* behavior.

Additionally, I believe that (because this dysfunction is ultimately rooted in our most social proto-emotion *Inclusion/Ostracization*) the onset of adolescence and the subsequent shift in brain chemistry that suddenly *heightens* the value of social inclusion in *non-kin* settings (the neural source of all that high school drama) essentially begins to give anorexia/bulimia & its dysfunctional circuitry *immense power* over behavior & decision-making at this point in life. In essence, the chemistry of adolescence is like a match that lights the tinder-keg of anorexia/bulimia's looming dysfunction.

This teen-aged emergence of the disorder's new power is mirrored by the ongoing construction & application of those modern behavioral systems: beliefs & rules (whose #1 priority is now *not eating*). And it's here that, according to our theory, we see the subtle-but-distinct differences emerge in how this neural dysfunction ultimately expresses itself: as anorexia or bulimia. There can be, obviously, a strong degree of overlap between those eating disorder sufferers who simply refuse to eat (anorexia or *restricting)* and those who sometimes eat, but regurgitate afterwards (bulimia or *bingeing & purging*). Eating disorder sufferers will often exhibit both behaviors to some degree. Nonetheless, research has shown that in addition to many individuals displaying only one or the other behavior, most "overlap" cases also show some clearly stronger tendency toward one behavior or the other. [28]

In the view of our theory, this distinction basically represents whether that individual's brain has come to favor a mainly *belief-based* or mainly *rule-based* strategy in pursuing their ultra-important *not eating* goal. Anorexia suggests a mainly *rule-based* strategy & bulimia suggests a mainly *belief-based* strategy.

The neural/behavioral difference between these two types of strategies mirrors the difference we described between *rule-based* ritualism & *belief-based* obsessive-compulsive disorder: a ritualist (anorexic) cognitively-but-unconsciously (& wrongly) *presumes* that they *must* perform this specific series of actions in order to achieve the intended (& usually specifically-defined) result of their ritualistic causal sequence. In contrast, a compulsive (bulimic) individual powerfully, consciously (& wrongly) *believes* that they *should* perform this (or these) action(s) because to *not* perform the action is—according to their belief—highly-likely to lead to an *ultimately* (& often broadly-defined or malleably-definable) bad result.

Basically, this means that anorexics' brains make it very difficult for them to engage in *any* bingeing, because their powerful rule-based behavioral sequences regarding *eating/not eating* simply do not allow for bingeing as part of the behavior. In contrast, bulimics' brains are generally more flexible in what they will allow—because they can set-up complex & interconnected beliefs that can *occasionally* permit certain behavior (*eating a lot* or bingeing) under the *self-promise* that it will be immediately followed by corrective behavior (*un-eating* or purging). This is a kind of belief-based rationalization.

A more rule-obsessive anorexic mind basically *cannot* "occasionally" permit *any* kind of food-related behavior—it *always* does everything almost *exactly* the same way in the pursuit of this unique *non-eating* goal. That's what rules are for: to be followed, always & without even thinking about it, *because they're rules*. As described earlier in this essay, the application of rules feels essentially inevitable & unconscious, while the application of beliefs feels like a conscious choice that we can make, and that we can sometimes convince ourselves to make a different one (*binge & purge* sometimes or *simply not eat* sometimes).

And in both anorexics *and* bulimics, these obsessively rewired & single-minded belief and/or rule systems can have a powerful impact on how the individual actually *perceives* (aka, *imagines*) their physical self—which can lead to the kind of body dysmorphia commonly associated with these disorders. No matter how thin you actually *are*, if your brain truly & powerfully consciously *believes* (or simply unconsciously *knows*) that the body it inhabits is "fat" then it will *perceive* the body it sees in the mirror as "fat" (and seek out any *actual* visual evidence that it can find to support this perception).

This disorder's high morbidity rate begs one question: how can we cure it? Our theory's full answer is longer than we have time for here (we've spent so long on this already that you've probably forgotten that explaining anorexia/bulimia isn't this essay's main purpose). But I will quickly say that the current model for intensive in-patient treatment (which frequently involves adhering to a long list of institutional rules & restrictions, and employs shaming/punishment-based strategies for enforcing those rules) is, unfortunately, a mostly wrong-headed approach.

What does a more ideal anorexia/bulimia treatment program look like? For one, it's done in an out-patient setting (developing new behaviors in a highly-non-real-world-&-isolated setting, then attempting to maintain those behaviors in an entirely different & highly new-stress environment is a perfect recipe for relapse & a significant waste of resources). During treatment, patients should live with relatives or other strongly-supportive (& healthy) individuals, but they should also have group meetings with other eating disorder patients & well-trained therapists on a daily basis at a local treatment facility (essentially, a "safe zone"). In addition, anorexics & bulimics should each receive specific treatment methods & therapy geared toward their different neural tendencies.

And maybe most importantly: there must be *no attempt at all* to employ *shaming* or *punishment-based* strategies as part of this behavioral therapy—to do so is as cruel (& dangerous) as throwing a burn victim into a fire. Tragically, the shaming of anorexics & bulimics (even those in treatment) is far too prevalent in our modern society, and it all-too-frequently has deadly results.

Moving on from the twin darkness of anorexia/bulimia —there are two other equally profound, yet not always devastating conditions that are (like psychopathy & OCD) the result of unique & seemingly-opposite neural circumstances (circumstances that are inborn, but are often expressed at different levels & with different developmental timing): autism, likely the result of overstimulated, indiscriminately-applied mirror neurons, and Asperger's, likely the result of non- or low-functioning mirror neurons. To categorize these unique neural circumstances as true deficits is, however, a mistake. I believe that, in truth, both of these "conditions" are merely another (and often an extraordinarily individual-specific) way to experience being.

And as shown by the myriad diverse & uniquely-talented individuals who possess these uncommon wirings, autism & Asperger's can also unleash the power of the human mind in surprising & amazing ways. Thanks again to the human brain's uncanny flexibility & its capacity to repurpose systems based on what other "normally" functioning mechanisms are available, "deficits" like the unique use of mirror-neurons can allow their highly-evolved power to be applied in those unexpected ways.

Consider, for example, the huge number of visual data points that those mirror-neurons are typically tracking when identifying, analyzing & remembering complex human facial expressions & physical movements. Now imagine that those resources are no longer devoted to human faces & movement, but used to help track & analyze *any* visually-composed palette. The repurposing of this power might help someone to, say, draw in detail the entire Manhattan skyline *from memory*—a skill demonstrated by the extraordinary autistic artist Stephen Wiltshire. Or imagine that those mirror-neurons allow an individual to internally, physically feel what it is to be a flag flapping in the wind *just by looking at a flag flapping in the wind*.

Or these unique autistic neural circumstances might produce someone like the legendary Temple Grandin—who applies her (often overstimulating) mirror-neurons & their empathy-producing capacities to the perspectives of other creatures (& combines this ability with that aforementioned enhanced ability to analyze complex visual patterns). These talents have allowed her to imagine & devise uniquely-humane & efficient slaughtering systems.

Of course, because mirror-neurons' typical highly-specialized facial analysis mechanisms help us to physically mimic the mouth movements required to first learn speech—severe autism can also disrupt speech development and, consequently, language acquisition. (The aforementioned artist, Wiltshire, was mute until the age of 5.) And interfering with language acquisition can powerfully alter the very *nature* of an individual's conscious experience. As with most systems in the human body, there's a very specific *give & take* involved with any unique circumstances.

Autism has also been shown to interfere with that "Theory of Mind" capacity we discussed earlier (predicting the internal experiences of others). This would make sense if, indeed, autism is primarily a result of atypical mirror neuron systems—because (as we hypothesized) our mirror neuron systems normally play a key role in intuitively understanding & predicting other people's feelings & intent.

In addition, those overstimulated and indiscriminately-applied mirror neurons can ultimately result in that commonly-observed *self-isolating* autistic behavior, which often includes repetitive physical acts or an intense focus on some external stimuli or pattern. The likely reason why these individuals seek such deep self-isolation is because their overactive & indiscriminate mirror neuron systems are *overloading* their pre-motor & somatosensory cortexes with all kinds of inappropriately-reflected incoming sensory data.

Indeed, it is hard for the rest of us to imagine what it might be like to experience the chaos of a young (& *barely-language-capable*) autistic mind as it is bombarded by powerful-but-disorganized sensory stimuli that is inappropriately reflected & experienced by those parts of our brain that help to define *our most innate physical perceptions of ourselves*. It is not hard to see how finding some excessively-repetitive set of actions or some deeply-immersive pattern to get lost within could provide exactly the kind of neural & physical relief that these people desperately seek: activity that might generate an intense (& reliably *predictable*) focus powerful enough to shut out the maddening & often *painful* chaos of the outside world.

This kind of autistic experience is also likely why strategies for pulling someone *out* of that chaos (& into a world that can be navigated) can be so widely varied: because basically *anything* that the brain can latch onto & use to begin making order out of the chaos can be the first step to "bootstrapping" this mind into a less chaotic world. However, because it's so difficult to have a decent idea what an autistic child's cross-wired systems might essentially *randomly* or *accidentally* latch onto, finding that path into their world (a path by which you might then begin to draw them out) can require almost *sleuth-like* observational skills.

Nonetheless, in all cases, the key to finding a path (if one even exists—for some the chaos may simply be too much) is truly attempting to *enter their world* and view their actions & desires (or *non-actions & non-desires*) from their point of view. If they enjoy something, *try to understand why*, and then *become part of their enjoyment*. The results of autism are ultimately expressed through deeply individualistic behaviors, and the best way to connect with any autistic individual is to truly share time in *their specific world*—to deeply ingrain yourself within the pleasure- & relief-seeking actions & patterns (& uniquely-individual *narratives*) that help to shape their experience.

In the case of Asperger's—where mirror-neurons are likely in opposite circumstances and *do not* effectively reflect *any* visual data to our somatosensory (tactile) & pre-motor (physical movement) cortexes—the absence of that chaos-generating *overstimulation* can make like life much more manageable than it is for someone with autism. In addition, many individuals with Asperger's also find ways to benefit from that neural-flexibility & repurposing. This means that if their mirror neurons aren't *reflecting* data, it seems that the brain can still often find a way to make analytical use of these powerful tools.

Thus, people with Asperger's tend to be better at organizing, associating & managing huge piles of *other* kinds of non-empathic data—like mathematic calculations or taxonomical information systems (exactly *how* these brain areas ultimately get repurposed likely depends on what new applications result in the most initial & ongoing pleasure, reward or relief).

Nonetheless, because developing children typically rely heavily on those mirror neuron's *reflective* capacities to help navigate social & person-to-person interactions, young people with Asperger's also have a strong tendency to exhibit their own types of (less intense) *self-isolating* types of behavior. And our typical reliance on those *reflective* capacities when learning complex physical actions is why individuals with Asperger's have more difficulty in honing such actions.

With Asperger's, these non/low-functioning mirror-neurons can also hinder early speech development, but again, it seems that these kinds of problems are typically much less severe than with autism. This is likely because the effects of autism are two-fold in regards to speech development & language acquisition: 1) specific, visually-perceived facial-data cannot be used to internally & physically mimic/learn speech acts, 2) all varieties of *non-human-focused* external sensory data are being reflected to the somatosensory & pre-motor cortexes, which actually *interferes* with the application of other systems in this speech-learning process.

In the case of Asperger's, this second problem is not an issue—which likely makes it easier for the brain to use other motor-script & rule-based (non-empathic) systems to aid in developing speech. The result is that early speech development efforts in these individuals are more deliberate, slower & less intuitive (essentially, less reflexive) than in typical neural circumstances, but once these motor scripts are learned and practiced, speech & language-use can still easily flower.

This lack of interference with compensatory strategies is likely why individuals with Asperger's typically have a much easier time than autistic individuals when integrating with social structures & circumstances geared towards "neuro-typical" individuals. Nonetheless, both Asperger's & autism can make it extremely difficult to empathically judge how others are feeling or behaving and thus, to respond appropriately. (And the subsequent lack of emotional data that these individuals reflexively derive from analyzing human faces also likely accounts for their natural indifference toward making eye contact when interacting with others.)

Which leaves us with our third category of causes for these deficits (or *uniquenesses*) of the mind: *system dysfunction from emotional or chemical imbalance*. This seems to be as inevitable in most people as the detrimental effects of aging.

We don't like to admit it, but almost all of us have some crossed emotional wires up there. It's hard for us not to—*knot* being the operative word. Because that's a lot what it's like up there in brain-town: one big spaghetti-bowl neural knot of data, emotions & associations. Our systems *mostly* handle that knot effortlessly, but stuff happens. Bad stuff. Sometimes it's bad stuff that ended up feeling way too good, reinforcing a self-destructive loop. Sometimes the bad stuff is tolerated because of an unnatural, overly-powerful fear of *even worse* stuff. In other words, *life is complicated*—and those complications can sometimes make our knot produce undesirable results.

This kind of *detrimentally-applied narrative logic* is at the root of much system dysfunction. These *are not* cases in which part of the system is physically misfiring, producing the kinds of chemical imbalances that lead to problems like bipolar disorder. These emotionally-based imbalances are actually a result of our systems doing exactly what they should be in response to the memory-stored & incoming data. But here the memory-data is producing some bad results. That's because the situation that led to that data was probably either emotionally extreme (like the highly-traumatic events of war) or painfully *twisted* (like being harmed by someone you love & trust).

In essence, these types of data use our own narrative-building & memory systems against our ultimate best interest. No matter how you slice it, this data is *trouble*. It's forcing us to use an *outlier* or a non-representative event to broadly shape our emotional responses counter-intuitively. *The war vet reacting violently to the tiniest provocation. The abused child growing into an adult who seeks an abusive spouse.*

In these cases, our view of the brain's memory tells us one very key thing about eliminating this kind of system dysfunction: *don't let the problem linger*. Because of the mechanisms of memory, every time these kinds of dysfunctional responses are triggered & repeated, the behavior only becomes more deeply ingrained and harder to change. When the brain begins to display this kind of dysfunction, it can quickly lead to a classic vicious cycle. Every almost-impossible-to-control behavioral response makes that response even more impossible-to-control in the future.

So how do we break that cycle? The key can be found right in those same memory systems. As discussed earlier, retelling a memory can slowly change the memory itself & its associations. This is partly our brain's way to keep narrative data up-to-date and optimally useful.

As noted, under normal circumstances repeated pinging (& retelling) is innate proof of data's usefulness. And if this useful data has been altered or embellished in the retelling, there has likely been a purpose: to somehow make the retelling *more* useful in that moment. The specific motivation for each alteration can vary. For example, some embellishments are intended to make the story more engaging for listeners. In other cases, some of the data may have degraded, and replacement data is inserted (a forgotten color detail replaced by a slightly different one in a retelling) in order to update the memory and keep it seemingly complete.

There's a shared secret behind all of these alterations: they somehow make us *feel* better when retelling the data. Engaging (essentially, providing pleasure to) listeners can evoke emotions like pride and generosity, and fixing a broken story detail likely gives us a little validity spike (essentially, more confidence in the story). And *sometimes* when a story makes us feel *bad* (like retelling a shameful act) we allow ourselves to change it little-by-little when we retell it, softening its sharp edges enough to make its retelling more tolerable.

Why would our brains let us do this? Because that painful narrative might contain some generally valuable data—after all, we do keep recalling it. Lessening the story's associated pain can allow us to use the data without having to suffer so much—which can lead to those unintended results, and may not be necessary anymore for our brain to retain the gist of the narrative.

This brings us back around to treating that trauma-spurred emotional imbalance. Psychologists often talk about the need to "process" bad memories in order to escape their self-destructive influence on behavior. This is essentially the above-described mechanism of changing a memory's emotional content & associations through retelling. Emotions are primarily narratively-produced, thus "reframing" the story when retelling and altering the narrative structure can help alter the emotions felt. These new emotions can now begin to help re-write that self-destructive memory data. Over time, if there's been enough change in the memory's emotional content & associations, pinging that memory no longer results in that bad data. We've "processed" the destructive memory using the mechanisms of our data-storage & narrative-building systems.

Frustratingly, these days pharmaceutical companies would like you to believe that this kind of emotionally-based system dysfunction is best treated by drugs. Let me be clear about my opinion here: *bullshit*. Drugs *might* be useful in some of the most extreme cases—helping to temporarily alleviate the most powerful, crippling emotional effects of the dysfunction in order to allow the memory re-writing mechanisms to do their trick. But even in these cases the drugs aren't really *solving* the problem, they're just helping to make it possible for the brain to use its own systems to solve it. Without engaging in talk or experiential therapy (like the highly-effective emerging virtual-reality techniques being used to treat PTSD) the drugs won't fix anything in the end.

In fact, I believe in almost all non-extreme cases of emotionally-based dysfunction, the drugs do more harm than good. There's no way to target a drug to one specific piece of memory data or a single set of narrative-building rules or a specific narrative-analyzing belief (which is *exactly* what we can do when "processing" or re-writing a memory). The drugs are making the *whole system* function improperly. Therefore, the mechanisms that need to do their jobs with precision in order to effectively rewrite that bad data are *also* being hindered by the drugs. This would seem to inherently make it harder for talk & experiential therapy to affect the necessary changes.

In the absence of one specific extremely unbalanced emotional response that *must* be mitigated for *anything* in the system to work effectively, the drugs likely do *almost nothing* to help solve the problem. You've simply numbed the whole system, and now your surgeon can't feel his fingers. Sure he was a little stressed & we thought calming him down would help his performance, *but not if he can't use his fingers*.

There are certainly those genuinely severe chemical imbalance or systematic emotional problems like bipolar disorder, schizophrenia, etc. (which we'll discuss in a bit) that require drug therapy to help make life manageable. Nonetheless, in the many *other* cases where some emotional imbalance is present, but not severe, I believe people are much better off training their *working* systems to compensate for these imbalances (instead of Zoloft, try *anger management*—a useful idea that has been given a bad name by practitioners who don't truly understand our emotional systems). The alternative is like taking Vicodin for muscle soreness—if you instead employ "hands-on" therapies and learn to live with it pharma-free, you'll likely be better at living.

Yes, there appears to be clear clinical evidence that a large percentage of individuals who suffer from symptoms like persistent, powerful & sometimes debilitating sadness or anxiety have demonstrated a lessening of these symptoms' persistence & intensity in response to drugs like anti-depressants. What I'm trying to point out here, however, is that using drug-regimens as a *primary* strategy for addressing these symptoms (particularly when the symptoms may be persistent, but not truly debilitating) is an inefficient & high cost approach to treatment—one that also has significantly fewer long-term benefits than a *neurally-rewiring* talk or experiential therapy approach.

The extra costs of these drug-regimens are both financial *and* neural. Financially, the amount of money that we *all* contribute (via insurance premiums & out of pocket costs) to the exploding profits of pharmaceutical companies is undoubtedly increased by the number of people who are nearly-automatically (& often-unnecessarily) prescribed some kind of anti-depressant immediately upon reporting symptoms. Neurally, the extra costs can come in the form of less sharp or fluid cognitive mechanisms, and generally less intense emotional experiences or responses. The problem with these costs is that they are neurally *global*. In other words, as described earlier, these drugs are not purposefully impacting the specific neural components that are actually the source of the symptoms—the drugs are impacting a global & fundamental mechanic that is broadly used *throughout* the brain.

The mechanic that the vast majority of these anti-depressants *globally* disturb is the management of serotonin in the brain (most anti-depressants fall into the category of either an *sri*, serotonin reuptake inhibitor, or an *ssri*, selective serotonin reuptake inhibitor). Can pharmaceutical serotonin management in the brain have a significant impact on the symptoms of depression & anxiety? *Of course it can*. Can pharmaceutical serotonin management in the brain also have a significant impact on a whole slew of other neural systems that play vital roles in maximizing our daily functionality & experience? *Of course it can*.

Are there alternative methods for alleviating these symptoms that do not challenge the effectiveness of all those other systems? *Absolutely* (e.g., that *neurally-rewiring* talk or experiential therapy.) Unfortunately, effectively applying those other therapeutic methods requires a deeper & truer understanding of our emotional & neural systems than many psychiatric professionals currently possess. Nonetheless, the *effective* application of other non-drug-based therapies can also have those additional *long-term* benefits that are not provided by most drug regimens (regimens that are basically built to keep an individual on the drugs for extended & often indefinite periods of time). The long-term benefit of *neurally-rewiring* your memories, rules, vocabulary & beliefs through therapy is that in *future* challenging emotional & cognitive circumstances, your brain will be *much better prepared* to effectively handle those challenges.

Messing with serotonin management in the brain for extended or indefinite periods of time ultimately has very few long-term benefits for the brain. In a way, the drugs simply put the actual systematic problem into "stasis"—still present in the wiring of the dysfunctional system, but its potency numbed by a general anesthetic. As soon as the anesthetic is removed, the still-wired problem is free to fully express itself again, which leads to a reapplication of the anesthetic, *etc.*, *etc.*, *etc.* This may be a wonderful business model, but it is clearly a poor treatment strategy.

And in many ways, most of the wide array of lesser "mood disorders" & similar diagnoses that the psychiatric establishment & pharmaceutical companies are misguidedly attempting to medicate into "normalcy" are phantom ailments. As we'll discuss near the conclusion of this essay, the human brain (like the rest of the human body) is purposefully designed (*aka*, evolutionarily-driven) to result in a variety of configurations, the vast majority of which are capable of effectively functioning within our world.

When your 3-year-old is projected to be shorter than 95% of the population, are you inclined to give them growth hormones? *I certainly hope not.* Why, then, are children whose brains tend to reward novelty & activity over deep engagement & sustained focus (aka, *ADHD*) medicated in order to achieve a more median level/type of mental engagement? Why are slight variations from the norm in brain traits less acceptable than slight variations from the norm in other physical traits? Have these other types of brains not proven to produce their own uniquely-useful results in previous human societies?

In fact, over the course of civilization, humankind's incremental progress has no doubt at times been powerfully *aided* by individuals whose brains possessed these more highly-varied & less conventional wirings. Consider that throughout history many of the most obsessive, hyperactive & risk-taking individuals have been among those who have pushed human exploration & discovery past existing boundaries (individuals who, in modern America, might be medicated into mediocrity before adolescence even arrives).

Yes, it's also fair to say that these categories of brains & the individuals they inhabit are more likely to find themselves at greater risk of personal harm (& even increase the risk-exposure of those closest to them). But the fact that these types of brains remain fairly commonplace in human society (much more commonplace than brains with *true* & *highly survival-adverse* disorders like schizophrenia) clearly indicates that the higher risk factors inherent in these less-conventional wirings has not outweighed the occasional benefit enough to result in a strong Darwinian *de-selection* of these neural traits among humans.

Quite the contrary: these brains seem to keep popping up in decent numbers because occasionally some of them can provide a few *awesome* benefits for the rest us of. Indeed, the rest of us might even tend to be more tolerant & *supportive* of such high-benefit (& possibly high-cost) eccentricity in order to continue reaping those broader gains—thus allowing this brain's unique wiring to aid in its *reproductively-benefitting* longer-term survival in a cleverly round-about fashion.

In other words, human brains aren't meant to be "perfect" (or to perform in exactly equal capacity & manner). They're meant to be *adaptable* & *malleable*—to both the specific needs of their environment *and* the needs of the social unit/structure in which they live. And humans do not build monolithic ant-like societal structures requiring nearly-identical parts that perform in exactly equal capacity & manner; our societies are complex & diverse structures that require a vast range of different brains & bodies to fulfill their various & multifaceted roles.

Ultimately, psychiatry's current determination of particular behavioral profiles as *dysfunctional* (those aforementioned & abundantly-diagnosed "mood" or "personality" disorders) is not founded upon any evolutionary or neural reasoning for defining them as "deficit-based" instead of simply natural & desirable variations within our adaptation-based species. It's merely that modern American society has both become more enamored with an everyone-should-be-normal-(&-*happy*) ideal and, at the same time, grown toward requiring a more monolithically-defined set of skills from its median & high-earning laborers, which has resulted in an educational system & culture that have also grown more monolithic in their goals (because no one seeks to be—or expends resources on cultivating—*low-earning* laborers, despite their absolute necessity within our society).

This has led a scientifically-unmoored & pharmaceutically-profit-driven psychiatric establishment to gear its own practices toward shaping individuals' behavior according to these newly-monolithic neural standards. And none of the aforementioned institutions has provided any sound reasoning for why their particular view of neural perfection ought to be considered the ideal model for *all human behavior.* It's simply that their model best fits the perceived needs of those humans who currently manage the economic & employment systems in America, humans who are primarily (& naturally) seeking to reap the most benefit from those systems for themselves & those around them (which is what humans & their ancestors have long been programmed to do).

Consider that there is not even a truly agreed-upon definition of the *symptoms* of many of these lesser "disorders" within the professional medical realm—and the psychiatric community fully admits that it has no truly biological basis for identifying many of these "disorders" or explaining why they are truly neural deficits instead of merely variations. But this certainly hasn't prevented these professionals from prescribing millions of pills in order to aid in "correcting" these mysterious, poorly-defined & often apparently completely-inexplicable "conditions." (Which, in the end, isn't much different from a carnival barker hawking neatly bottled & labeled "remedies" from the back of his horse-drawn wagon.)

In our early 21st-century America, probably the most egregiously erroneous & damaging of these phantom diagnoses is that aforementioned & quickly-becoming-infamous acronym: *ADHD* (Attention Deficit Hyperactivity Disorder). Let me also be clear about my opinion on this: *ADHD is B-U-double-hockey-stick-Shit*. Yes, some kids are more hyper than others, and they also tend to be more easily distracted, making it harder for them to sustain focus. *So what.*

This is not a disorder. It is merely a human brain that's developing along a less-common, but still functional & useful path. We've simply taken the biggest bulge in the bell curve, decided it was easier to use a one-size-fits-all educational/behavioral strategy, then declared everyone outside the bulge *dysfunctional*, and we're now trying to medicate them back into the bulge with dangerous stuff like Adderall —more commonly known as *speed*. (And when *truly uncontrollable* behavior in children is tagged as ADHD, that's just pure *misdiagnosis* of an *actual* neural problem.)

Of course, what's really happening (and was eminently predictable) is that we're turning lots of slightly-outside-the-bulge kids into *speed addicts*. And we're so cavalier about applying this phantom diagnosis that we're prescribing *even more* speed to a bunch of teens & young adults who are *pretending* to be slightly-outside-the-bulge just so they can, y'know, *take some speed*—which has lots of very short-lived, but very awesome benefits that result in lots of long-term problems. (And young people's brains are naturally totally enamored by those kinds of emotional equations.) Once again, like most of those prescription-triggering lesser "mood disorders,"*ADHD* is not even a neurally-defined phenomenon. It is merely & flimsily a vague, scientifically-baseless set of "diagnostic" standards. *ADHD* is, essentially, the result of a *questionnaire*—one whose imprecision does not impede its power to recommend pharmaceutical remedies.

Even deeply research-based examinations of newly-emerging hypotheses for ADHD's neural basis—like the insightful 2010 paper "*Is the ADHD brain wired differently? A review* *on structural and functional connectivity in attention deficit hyperactivity disorder*" by German neuroscientists Kerstin Konrad & Simon B. Eickhoff [29]—mostly conclude with some academic version of: *basically, beyond some broad & contradictory strokes, we still have no idea what actually causes this or what it's really all about.*

One exception is a recent 2017 brain-imaging study that claims to substantiate ADHD's status as an *actual* neural dysfunction by identifying multiple subcortical brain regions that appear *underdeveloped* (display below-average volume) in individuals who present ADHD's "symptoms." [30] This is, in our view, completely *unsurprising*—because we merely see these *less-common* (outside-the-bulge) behavioral profiles to be a simple case of *less-common* (but functional) neural developmental arcs, ones that are no more "dysfunctional" than those less-common *physical* developmental arcs. And our view is bolstered by that same study's observation that many of these individuals' *underdeveloped* brain regions eventually "catch-up" to "normal" brains by adulthood. In other words, the study presents no real proof that ADHD is the result of some neural *disorder*. Rather, it seems to support the notion that these brains are merely developing along different, less-common, but *within-normal-variance* arcs.

Considering all of the broad, deep uncertainty surrounding ADHD, the over-diagnosis of this phantom condition (and the resulting over-prescription of speed to children) represents nothing less than an epidemic of malpractice —an entirely unnecessary one.

There is, however, one emotionally-based neural situation that is not a true neural *deficit*, yet ought to be seen as its own disorder: sociopathic behavior. As we mentioned when discussing its difference from psychopathic behavior, according to our theory sociopaths actually possess fully-functioning neural systems. The disorder results from learned & highly-destructive—either self-destructive, societally-destructive, or both—rules & beliefs that compel & allow this person to act violently or callously in the service of achieving their goals.

Most repeatedly-violent criminals are essentially, at some level, sociopaths. And I believe that reforming these individuals' belief & rule systems in a way that makes them less destructive in society is harder than is typically assumed (and for the most part, is not typically achieved nor seriously attempted in the American penal system). This is because it's likely that the best strategy for reforming these sociopaths is years of intense & individually-tailored psychotherapy administered by a well-trained expert—combined with a living environment that powerfully rewards socially-constructive behavior & provides strong models of such behavior. I don't think I'm going out on a limb by saying that our penal system does not employ these strategies—nor does it have any intention to.

Moving on to those primarily-emotional disorders that actually *are* the result of genuinely severe chemical imbalance or neural-system problems—there are two that appear, like autism & Asperger's, to be opposite dysfunctions in essentially the same system: bipolar disorder & major depressive disorder.

In bipolar individuals, those core pain & pleasure emotional poles at the root of *all* emotions seem to be prone to drastic swings, resulting in the extreme expression of emotions located within the currently-dominant pole. Theoretically, this could be a result of something like dysfunction in the output of left-brain emotional equations (i.e., a processing glitch that always leads to "maximum" pain- or pleasure-based responses) or dysfunction within the mechanisms that produce neurotransmitters as a result of those emotional equations. In either case, the result of bipolarism is an individual whose "average" emotional responses occur at levels that actually far *exceed* a typical average—leading to swings between mania (hyper-positivity) and deep sadness (hyper-negativity).

Major depressive disorder is often mistakenly equated with the kind of deep sadness that is experienced during the downswings of bipolarism. But actual accounts of the experience from major depressives suggests something much different. Depressives actually typically describe the worst part of their experience as the *total absence of any kind of feeling*—positive or negative.

One way to view the primary difference between bipolar disorder & major depressive disorder: the former is essentially *too much* intense emotion, while the latter is essentially *no emotion at all*. In the view of Narrative Complexity, major depressive disorder thus suggests a problem with a mechanism that we mentioned back in Essay #2 after the Emotions Matrix: our *engagement/boredom* mechanisms. In our model, these emotional mechanisms would be impacted by data judgments like *novelty* & *relevance*—helping to determine if the incoming data is particularly unique or useful (positive novelty & relevance judgements= engagement, which produces actual emotions; negative novelty & relevance judgements=boredom, an emotionlessness that leads us to seek something else to be engaged by).

In major depressives, it seems that this engagement mechanism simply doesn't *engage*—making any subsequent emotional production essentially impossible. In other words, these individuals are perpetually & soul-draining-ly *bored*—utterly craving some stimuli or interaction that might result in some actual emotion (which is a craving that boredom is *meant* to trigger). And despite this empty craving, they can't even really *imagine* feeling anything again. It's a kind of bottomless hollowness.

Thus, according to our theory, this appears to be a specific dysfunction in the emotional mechanisms that *employ* those novelty & relevance judgements. Essentially, for someone with major depression: nothing *interests* them & nothing *matters* to them (aka, nothing is *emotionally* responded to as novel or relevant).

And on the blog "Hyperbole And A Half" I found a fascinating, personal, first-hand anecdotal description of a depressive episode in which the experience that triggered their suddenly-growing emergence from utter boredom was oddly (& almost solely) *novelty-based*: the sight of a single, lonely piece of shriveled corn lying astray beneath the refrigerator.

In this account, the inexplicable, but palpable *oddness* of this sight, and the strange way in which it somehow perfectly symbolized this individual's lost state of being—the quirky *connection* between a highly-novel judgement & a personal judgement that it closely matched (aka, *high novelty + strong relevance*)—this experience suddenly set off a powerful & outsized attack of hysterical, uncontrollable *laughter.* (And in our essays' *Comedic Addendum*, we explain the vital role that novelty plays in humor.)

This burst of powerful novelty-based emotion began parting the clouds—helping their deeply-depressed-self into the world of the *feeling* again. It's as if this connection finally brought the brain's novelty & relevance responses back online, allowing it to emotionally *engage* once more. This would, indeed, make sense if major depressive disorder was essentially a dysfunction in our novelty/relevance-employing emotional mechanisms—a dysfunction that prevents the *engagement* required to produce *any* emotions.

The last major systematic neurally-based dysfunction we'll discuss is the one that remains most mysterious in the view of our theory: schizophrenia. One of the factors that makes schizophrenia so confounding is that is seems to be both a broadly-based & a traveling-over-time neural problem, resulting in behavioral dysfunction that can be expressed differently as individuals age. Studies have suggested that in people with schizophrenia there are often specific portions of the brain that exhibit a loss of gray matter—basically, these areas possess less functional neural tissue than normal. [31] This problem appears to begin in parietal lobe regions that support visuospatial and associative thinking. As it progresses, the problem seems to reach more high-functioning & perceptual areas of the brain —leading to more severe psychosis, i.e. powerful hallucinations, and false-but-convincing narratives that can prompt (outwardly-nonsensically) obsessive or anxiety-ridden behavior.

In an earlier version of this book, we theorized that the “traveling” tissue death observed in schizophrenics suggested that the problem could be, at its root, vascular. (And recent brain research has detected evidence of vascular dysfunction within the brains of schizophrenics.[32, 33]) However, a study published in January of 2016 has shed new light on schizophrenia’s possible genetic roots.[34] The research provides compelling evidence that schizophrenics’ neural dysfunction is a result of “overly-aggressive” synaptic pruning (a mechanic that is a vital to brain development). And as time passes, the cumulative damage from too much synaptic pruning increases the severity of the neural dysfunction.

Regardless of how the dysfunction ultimately spreads, this broad range of shifting symptoms seems to make it very possible that much diagnosed schizophrenia is actually other as-yet-unidentified brain dysfunctions that present similarly to one stage or another of schizophrenia and are conveniently tossed into the schizophrenia basket. There is much research to be done before we can definitively identify the full causes & pathology of “true” schizophrenia, but a good starting point might be to require the observation of *multiple* dysfunctions that change or progress over time (typically creating greater-over-time interference with conscious perception).

In other words, currently the sudden appearance of auditory hallucinations (which can lead to all kinds of other symptoms) in a 25-year-old might be diagnosed as schizophrenia (despite no previous emotional or mental problems) when it is, in fact, a specific problem in the auditory cortex (which can lead to other processing problems that result in those other symptoms). As we've noted before, until we can get a good look at what's actually happening "under the hood" there are likely to be many unique neural circumstances & disorders that will remain a mystery at their root.

A Final Mantra: Don't Lose Your Mind

I was a teenager of the 80s. Thus, vividly imprinted into my memory is the image on the cover of Douglas Adams' book "The Hitchhiker's Guide to the Galaxy"—an iconic circle & “thumbs-up” above the words *Don't Panic.*[35] In addition to being a delicious wordplay joke (the icon suggesting a *Don't-Panic* “button”) it was, for me, a little subversive mantra. A way to remind myself when necessary to *stay in the moment*. If I were to create a subversive little mantra for this essay, it would be: *Don't lose your mind.*

The wordplay here is a bit more dire than Adams'—our hidden image is not a Wile E. Coyote-esque adornment, but rather, an amnesia-induced nightmare. Nonetheless, its meta-message is the same: *be in the moment.* If there's one central life-lesson that we can learn from our exploration of the brain's data storage & handling systems, it's the value of being *present*—attentive & engaged—in our moment-to-moment experiences. Maintaining this strong, immersive & interactive connection with our environment & its people is the best way to *optimize* those magnificent systems of consciousness. These are the moments that make us *feel* and *remember* our lives—that help us to powerfully sear our neurons with those vivid memories whose stories & detail create that deep, rich texture of a fully-lived existence. *(All of which is a strong argument against living your life through a screen in your hand.)*

Conversely when we say "*Don't lose your mind*" we're also reminding ourselves to *be* in *all* moments (or as many as possible) present *and* past. More succinctly, *remember.* That is, after all, the real sum of our mind: the accumulation & ongoing recollection of all those moments in which you were once *present* enough to create a memory. As evidenced in heartbreaking examples like Alzheimer's, once we lose access to all of that remembered data, we truly do lose our minds—which is, of course, to lose our *selves.*

Who are you? In a strange (but real) way you are a constantly-shifting location in that ever-humming cerebral cortex. Who we are in any given moment is essentially comprised of what we have *access to* within our data storage at that moment. And amongst that lumpy, folded, gray matter, our access to all of those other moments comes through the moment we are *currently* occupying.

This potential-memory thought-parcel provides each moment's doorway to everything you currently are. That fresh eddy of neurons—alive with new energy & associations—provides the propulsive force that catapults our mind's lighting both forward in time—to our next thought—and backward through our history, into the modules of our memories. This *at-the-moment* location in your neurons and its capability through association & construction to bring forth all the necessary information to know & produce *who you are* is—in that strange-but-real way—where *you* centrally *exist* at this very second. *Don't lose your mind.*

If we looked at ourselves this way more often, we would likely take better care of this almost-magical machinery inside our skulls. *(Again, why you should always wear a helmet.)* And not just protect it from physical harm, but work diligently to keep it active & robust, to avoid feeding it a steady stream of narrative junk food that's all short-term pleasure with no long-term gain (which is why someday you'll *barely remember* most of the junk), to seriously consider the *emotional impact* of how we behave within, respond to & *think about* our lives. All of these are key factors in how our brain remembers, associates & constructs the self-building data in our minds. These are the things that make us *who we are*.

At the heart of all this is that aforementioned conundrum: *nature vs. nurture*. There are plenty of people who would have you believe that one or the other has the upper hand in this "battle." But in truth it's not a battle at all, it's a *joint engagement*. Our brains are designed to allow nature & nurture to work together in building who we are in a way that best adapts our particular genetic expression of a human to its specific environment. This is why humans are so *awesome*. We're like these amazingly-differentiated Lego pieces intended to work complexly & interchangeably together—creating systems & structures (of all kinds) that ultimately aid in the propagation of more human genes.

Our variable *natures* are determined by subsets of subsets of systems within all of us that each have slightly different levels of inborn functionality. In a big-system physical way we can see this expressed in people with different visual acuity, muscular strength, height, lung capacity—*every* system allows for lots of variability that does not overtly harm overall functionality. This aids the human genome in producing all of those awesomely-interchangeable Lego pieces, which in turn allows human society to fill all the various roles needed to maintain & build its complex systems & structures—protectors, thinkers, crafters, cultivators, and on and on. Humans are, in essence, a bizarrely macro version of those modular neural components—an externalized societal expression of that highly-adaptable system of programmable, interchangeable, malleable, associated parts that is our network of neurons.

And our variable inborn nature is also powerfully expressed in the systems that *construct* that neural network. Certain brains release certain chemicals in slightly different increments; others handle the results of certain emotional equations in ways that produce slightly different data outputs; others exhibit a greater natural fluidity in certain synaptic structures; others possess a slightly more robust capacity to match multiple patterns. Within our deeply-interwoven systems, such variations can produce a vast array of different types of human brains. And all of our own individual variations makes each of us more prone to make certain kinds of choices in certain kinds of situations. This is the *I am who I am* part of our minds—the very-hard-to-change tendencies of personality that continually shape our path through existence.

Why would human evolution allow brain functionality that is so highly variable? In other words: *Wouldn't it be better for everyone to be as "smart" as possible? Wouldn't humans with the most-brilliant processing systems have been most likely to survive our evolution?* Not necessarily. This is because everything in brain development is a *trade-off*. This is obvious in our main physical attributes (e.g., more brute strength is likely to lessen speed & agility). And those brains with greater processing speed & a powerful capacity for more complex pattern analysis & construction—they seem to be (we'll try to be delicate here) more likely to become *unstable*. (An analogy: *when something has more strands, it's easier for things to get tangled.*)

Thus—like everything evolution seems to spit out—the variability of human brain capabilities appears to be the result of a risk/reward proposition. As a species, we're better off in general if we employ both powerful, unstable processors and less robust, but more-reliable machines in our mix. In addition, these different types of brains would be likely to ultimately *desire* different kinds of roles in society, aiding in that world-winning evolutionary strategy of *intra-species Lego-ization.*

The other partner in this *joint engagement*, however, is an equal titan in the matter of self-building. *Nurture* is no weakling. Although our nature is responsible for our innate tendencies to behave in certain ways, ultimately the actual decisions themselves are primarily determined by all that data we've *experienced*—by our memories. This is the *you are what you eat* part of our minds.

The way in which those memories impact how we view ourselves and the choices we make—the very *lives we live*—has been the topic of this entire essay. In essence, it's the topic of all these essays. Because whether we're talking about our emotions, dreams, or internal dialogue loop, all of those mechanisms of consciousness require one primary element to make them run: data. And sometime around the age of 2 or 3 the vast majority of the data that influences *who we are* comes not from the present moment, but from our data storage banks, our *memories*.

Don't lose your mind. I actually have some personal experience in this mind-losing area that helped spark many of the insights in these essays. As mentioned in the first essay, I had my own (and thankfully brief) Alzheimer's-esque episode in my late 30s that produced myriad fascinating results. In short, I had a migraine-induced seizure that resulted in an hour of unconsciousness & subsequent temporary partial-amnesia (and included that brief sensory-deprived, but linguistically-conscious experience described in Essay #1). The most severe amnesia lasted only a couple weeks, and within 3-4 months my mind was mostly back to maximum efficiency—although "short-term" memory problems persisted for a little longer.

And another *odd effect* that persisted: weak validity judgements of old data. e.g., I could fully & accurately write-out from memory a frequently-used, but 20-year-old pancake recipe—yet when I looked at my written-out recipe on paper, I couldn't actually tell whether or not it was, indeed, accurate. For months, I had to call my mother to confirm the recipe every time I made the pancakes. It was weird—but may have been the result of the long-ingrained recipe being present in (& executable via) my rules resource, yet still *unconfirmable* via the hard-to-access memory-based recipe data (whose incoming paths had been buried by my seizure's neural electrical avalanche).

When I first awoke in the emergency room, I knew, generally speaking, *who* I was, but not much else. My memory of that moment (and yes, I *do* ironically have a memory of a moment in which I essentially had no memory) is primarily one of feeling lost & *embarrassed*—embarrassed that I didn't know anything about how I'd gotten there or where I even was in the course of my own life.

And the experience that I had over those next few hours was the beginning of the fascinating strangeness of forgetting, of *losing your mind*. Only one thing concerned me after I awoke: building a story about myself. At the time we were expecting our second child, I was applying (unsuccessfully) for a fellowship, and it was (very importantly) the middle of a Bears' season. And as I emerged from my fog, those were the narrative threads that I felt compelled to fill in. Totally lost in time, I asked over & over: *Did we have the baby? Did I submit the application? Are the Bears having a good season?* (I'd watched the team get pummeled by the Cardinals earlier that day, and my Bears obsession follows me everywhere: once under the effects of the anesthetic *versed*, the only thing I wanted to talk about was an injury to defensive end Alonzo Spellman's shoulder.)

I asked those questions repeatedly because, although the vague idea of each was stuck in my head, I had no narrative within which to place them. And my brain *needed* narratives desperately at that moment—something to hang my hat on, to help me say something more to myself than the bare-minimum *I am here*. Simply speaking, my consciousness wanted to do its job. But without proper access to its data-banks, the narrative-building machine was sputtering & coughing out confused nonsense. And it *recognized* this pattern-less data as nonsense, so it kept putting questions into the prompt, seeking the information it needed to complete its equations and get the loop flowing again.

The only information it had at its disposal was the most-basic, strongly-imprinted, self-defining data—conveniently stored in that left-brain vocabulary resource. Although my right-brain-based memories would take weeks to become truly functional, that vocabulary resource (as well as other narrative-building left-brain mechanisms) seemed to come online fairly quickly & smoothly; the evidence of this being my reasonable (although still fuzzy) ability to understand language, talk, identify people, and answer the question: *Who are you?* (Just thinking your name or seeing your face in a mirror sets-off a cascade of super-imprinted self-definitions: *I'm a writer, Rebecca is my wife, Vivienne & Camellia are my daughters…*)

The only other data that kept popping up was those vague ideas—*the baby, the Bears, the application*—whose presence was the likely result of their recently frequently-recalled nature & high-priority. It's the same as the reason why our dreams use this kind of data to start their dialogue loop: it's the most handy & available, *right here on the shelf*. But it wasn't until my access to that larger data-bank began to grow that I was able to begin feeling like *myself*. As soon as I was able to start building narratives about the present with the help of stored data, I was able to get my bearings.

And one of the most interesting things about the next several weeks—as the severest amnesia faded—was how what I *remembered* all seemed to come in narrative strands. In other words, it wasn't like my memory slowly & sequentially expanded deeper into the past—first remembering last week, then last month, etc. Rather, certain narratives suddenly became available. "*Right, we took Vivienne out to Fairfax for Halloween, and we…*" or "*I had finished that part of the poem, but I was going to change…*"

In the latter case, when this narrative became available I suddenly remembered very specific details about the revisions I was intending (and even a kind of *nuance* about my ambivalence over making them). I'd been putting off the revisions and hadn't thought about them in awhile, so that data was actually several weeks old—yet some of the more recent memories did not return until later. And some memories never returned at all. Although, as a frequent journal-writer, I still had some record of these memories. Yet, to this day, when I reread those unremembered entries, it *feels* like they were written by someone else.

There was no temporal pattern to my re-remembering. The remembering occurred *narrative-by-narrative*. And I can say from experience, without access to those narratives, you feel exactly how we might imagine a ghost does: *here, but not*—temporally displaced & terrifyingly at sea. In these moments, we are afraid of one thing above all others: that we might never return. *Don't lose your mind.*

In the end, our brain is designed to function as (and, if necessary, rebuild its architecture from) only the minimum version of ourselves—the *I am here now* version who is usually the first to reappear from any particular neural ether (this is even who we often are momentarily when we first awake in the morning). And although this *simple being* may be good enough for our brains, it's not good enough for *us*. We are, quite reasonably, desperately attached to all of the data that we've grown so familiar with & dependent upon. To feel it vanish is nothing less than the purest & most profound sense of loss that we can imagine.

During that decade-ago afternoon with my grandmother, just as we were leaving my uncle's house, a moment occurred—it was the moment that I feel is last time I ever really saw *my grandmother*. I was in the doorway, saying goodbye to her, when suddenly from nowhere she *returned*. You could see it in her eyes, that human thing that knows itself & its place in the world had come to life one more time. This was not the temporally displaced ghost, it was my *grandma*.

Although the confusion was still there, I could see how she felt in the way she looked at us (partly through that powerful tool—empathy). I could see what she knew: *that we were all together for the holidays, that we were her grandchildren, that we were leaving her.*

And there was one other piece of knowledge I could see in her mind, something that was likely accessible because it had long been a dominant thought: *I am sick & my memory is fading*. This knowledge made her say goodbye in a way that she knew she might never be able to again. And the tears our goodbyes produced in her, the sadness of the moment she actually *perceived* came from one basic narrative that her mind could still process in the moment: *"I am here with you now, and I may never be again, because I am losing my mind."* She is—as I write this—still *here*, but her *stories* are not.

This is how we know ourselves, how I rebuilt my *self* from the ether—by assembling the only pieces of *being* that we can find lying around our neurons: *narrative-by-narrative*. Like my grandmother's rote recitation—*word-for-word*, *smile-for-smile*—when all that's left are the barest-bones of our mind to view amidst the emptiness, the only things we can truly see are the *stories* in those bones. These are the supple skeletons of words & syntax that carry with them the flesh of our lives. *Don't lose your mind*. In the end, this can only mean one thing really: hold tight to your stories, your memories—*they* are who *you* are.

###

Endnotes:

p.102:
1. *Kitamura, Takashi, et al. "Engrams and circuits crucial for systems consolidation of a memory." Science 356.6333 (2017): 73-78.*

p.103:
2. Edelman, Gerald M., and Giulio Tononi. *A universe of consciousness: How matter becomes imagination.* Basic Books, 2000.

3. Deacon, Terrence. *Incomplete Nature: How Mind Emerged from Matter*. Norton, 2011.

4. Tse, Peter Ulric. *The Neural Basis of Free Will: Criterial Causation*. MIT Press, 2013.

5. Hofstadter, Douglas R. *I Am a Strange Loop*. Basic books, 2007.

6. Kilpatrick, Lisa, and Larry Cahill. "Amygdala modulation of parahippocampal and frontal regions during emotionally influenced memory storage." *Neuroimage* 20.4 (2003): 2091-2099.

7. Lenz, F. A., et al. "The sensory-limbic model of pain memory: connections from thalamus to the limbic system mediate the learned component of the affective dimension of pain." *Pain Forum*. Vol. 6. No. 1. Churchill Livingstone, 1997.

p.107:
8. Hafting, Torkel, et al. "Microstructure of a spatial map in the entorhinal cortex." *Nature* 436.7052 (2005): 801-806.

p.110:
9. Ford, Judith M., et al. "Cortical responsiveness during inner speech in schizophrenia: an event-related potential study." *American Journal of Psychiatry* 158.11 (2001): 1914-1916.

p.123:
10. Spencer, Susan S. "Corpus callosum section and other disconnection procedures for medically intractable epilepsy." *Epilepsia* 29.s2 (1988): S85-S99.

p.125:
11. Halliday, M.A.K., and Matthiessen ,Christian M.I.M. *Construing Experience Through Meaning: A Language-Based Approach to Cognition.* Continuum, 1999.

p.130:
12. Deacon, Terrence. *The Symbolic Species*. Norton, 1999.

p.132:
13. Robertson, Brita, et al. "GABA distribution in lamprey is phylogenetically conserved." *Journal of Comparative Neurology* 503.1 (2007): 47-63.

14. Pombal, M. A. "Afferent connections of the optic tectum in lampreys: an experimental study." *Brain, behavior and evolution* 69.1 (2006): 37-68.

15. Burgess, Neil, Eleanor A. Maguire, and John O'Keefe. "The human hippocampus and spatial and episodic memory." *Neuron* 35.4 (2002): 625-641.

p.133:
16. Fyhn, Marianne, et al. "Spatial representation in the entorhinal cortex." *Science* 305.5688 (2004): 1258-1264.

17. Zola-Morgan, Stuart, Larry R. Squire, and D. G. Amaral. "Human amnesia and the medial temporal region: enduring memory impairment following a bilateral lesion limited to field CA1 of the hippocampus." *The Journal of Neuroscience* 6.10 (1986): 2950-2967.

18. Jacobs, Joshua, et al. "Direct recordings of grid-like neuronal activity in human spatial navigation." *Nature neuroscience* 16.9 (2013): 1188-1190.

p.134:
19. Ullman, Michael T., et al. "A neural dissociation within language: Evidence that the mental dictionary is part of declarative memory, and that grammatical rules are processed by the procedural system." *Journal of cognitive neuroscience* 9.2 (1997): 266-276.

p.140:
20. Marsolek, Chad J., et al. "Hemispheric asymmetries in motivation neurally dissociate self-description processes." *Emotion* 13.3 (2013): 462.

p.152:
21. Oberman, Lindsay M., et al. "EEG evidence for mirror neuron dysfunction in autism spectrum disorders." *Cognitive Brain Research* 24.2 (2005): 190-198.

p.157:
22. Jaeggi, Susanne M., et al. "Improving fluid intelligence with training on working memory." *Proceedings of the National Academy of Sciences* 105.19 (2008): 6829-6833.

p.160:
23. Zeman, Adam, Michaela Dewar, and Sergio Della Sala. "Lives without imagery–Congenital aphantasia." *Cortex* (2015).

p.163:
24. Graziano, Michael SA, and Taylor W. Webb. "The attention schema theory: a mechanistic account of subjective awareness." *Frontiers in psychology* 6 (2015).

p.170:
25. Sacks, Oliver. *The man who mistook his wife for a hat: And other clinical tales*. Simon & Schuster, 1998.

p.172:
26. Mather, Mara. "The emotion paradox in the aging brain." *Annals of the New York Academy of Sciences* 1251.1 (2012): 33-49.

p.175:
27. Kiehl, Kent A. "A cognitive neuroscience perspective on psychopathy: evidence for paralimbic system dysfunction." *Psychiatry research* 142.2 (2006): 107-128.

p.178:
28. Kaye, Walter. "Neurobiology of anorexia and bulimia nervosa." *Physiology & Behavior* 94.1 (2008): 121-135.

p.191:
29. Konrad, Kerstin, and Simon B. Eickhoff. "Is the ADHD brain wired differently? A review on structural and functional connectivity in attention deficit hyperactivity disorder." *Human brain mapping* 31.6 (2010): 904-916.

30. Hoogman, Martine, et al. "Subcortical brain volume differences in participants with attention deficit hyperactivity disorder in children and adults: a cross-sectional mega-analysis." *The Lancet Psychiatry* 4.4 (2017): 310-319.

p.193:
31. Thompson, Paul M., et al. "Mapping adolescent brain change reveals dynamic wave of accelerated gray matter loss in very early-onset schizophrenia." *Proceedings of the National Academy of Sciences* 98.20 (2001): 11650-11655.

p.194:
32. Keefe, Richard SE, et al. "Microvascular Abnormality in Schizophrenia as Shown by Retinal Imaging.

33. Hanson, Daniel R., and Irving I. Gottesman. "Theories of schizophrenia: a genetic-inflammatory-vascular synthesis." *BMC Medical Genetics* 6.1 (2005): 7.

34. Sekar, Aswin, et al. "Schizophrenia risk from complex variation of complement component 4." *Nature* (2016).

35. Adams, Douglas. *The Hitchhiker's Guide to the Galaxy*. Harmony Books, 1980.

The Will Of The Free

After all the vessels of philosophy have been coaxed into the harbor, after all the boats carrying all the theories of mind & their mechanisms have been tied snugly to the piers of science, one vessel seems always to remain stubbornly at sea—and painted on its bow are two words: *Free Will*.

Do humans possess free will? There are a plethora of ways to approach this question, but most can be divided into two categories: philosophical & scientific. We might think of the philosophical approach as attempting to answer the question: *what truly is free will?* In contrast to that broader query, the scientific approach attempts to discern the specific relationship between the neural activity that determines our actions and the neural activity that generates our conscious contemplation of those actions. In other words, science attempts to answer the chicken or egg question: *which comes first, thinking about doing it or deciding to do it?*

In this last essay, we'll use the currents from both of these approaches to finally bring our wayward vessel, *Free Will*, into harbor alongside the rest of Narrative Complexity's multi-faceted fleet. But we'll begin by exploring that more concrete question about the neural relationship between thinking & deciding; this is the question that's most directly addressed by the elements & mechanisms of our theory here. And since free will is essentially about whether we are consciously capable of "choosing" our actions, at the center of this question is the relationship between our *conscious mind* & our *unconscious (or subconscious) mind.*

One Brain, Two Minds

How, then, do we define these two entities of mind? In the view of Narrative Complexity, our *conscious mind* is essentially that *experience* of hearing the internal dialogue generated by our loop of consciousness. The locus of this part of our loop (our *Dynamic Core* of conscious awareness) appears to be located within humans' highly-advanced prefrontal cortex.

Thus, our actual consciously-experienced mind is merely a *portion* of that loop—the narrative parcel that finally arrives on our conscious highway of internal dialogue & becomes one of those candidates for a short- or long-term memory. This experience includes our "awareness" of the current moment (or of those surrounding elements of the moment that our conscious focus has prioritized & is actively attending to).

Some of the sensory, internal & emotional data from that conscious "awareness" can be attached to those narrative parcels as they are seared in our storage, creating a more full memory of that experience—a memory that is a combination of internal dialogue narrative parcels & associated spike experiential data that occurred simultaneously. Thus—when we recall those right-brain-based narrative memories—what we are recalling was originally (in its *very first* iteration) founded upon a low-fidelity, attention-defined recording of our *conscious mind* in those moments.

When we compare this view of the conscious mind with the many mechanisms in our system that appear to occur in the loop *well before* that conscious dialogue actually arrives in our heads, we can get a good idea of just *how much* of our consciousness-generating mechanisms are actually occurring within the realm that we would define as the *unconscious mind.* In fact, you might even say that Narrative Complexity is as much a theory of *unconsciousness* as it is a theory of consciousness.

In truth, thus far we have left the majority of the unconscious mind out of the discussion, and have mostly focused on those elements of mind that contribute directly to our conscious experience—like emotions & internal dialogue. Luckily for us, there's no need for our theory to present its own detailed explanation of the mechanisms of the unconscious mind, because in 2008, Yale University's John A. Bargh & Ezequiel Morsella published a new & compelling view of the unconscious mind[1] whose perspective aligns perfectly with our theory's approach to the matter. (And Morsella's own insightful *Passive Frame Theory*[2] presents a useful framework for understanding why the *conscious* arena has evolved specifically to help direct the kinds of voluntary, adaptive skeletal-muscular actions that we'll be discussing.)

The basic thrust of Bargh & Morsella's argument is that the currently dominant view of the unconscious—essentially equating it with the "subliminal"—provides far too narrow a perspective on its genuinely robust systems. In many ways, most current discussions of our subconscious treat it as an undercurrent of thought that subtly & unknowingly influences our much more dominant conscious faculties. Bargh & Morsella argue that in reality, our unconscious mind is the primary (& ancient) machinery of the brain, and our conscious systems are actually a final layer of experience set atop that machinery. This is a perspective that Narrative Complexity wholeheartedly shares. (It's also shared by Gerald Edelman, the author of the *Dynamic Core Hypothesis*[3] that our theory relies heavily on.)

This Old Hominin

Here's one way to think of it: roll back the evolutionary clock to the time before humans developed self-aware consciousness—back to those brains that had not yet developed the capacity for internal dialogue. Since we know that even the most advanced apes do not appear to possess the full capacities of human internal dialogue, we might assume that those virgin branches of our hominin ancestry likely arrived on the evolutionary scene without those full capacities having yet been developed. (An assumption we explored back in essay #1 while discussing our *Terrence Deacon-supported* [4] theory of language evolution.)

How might we think of these ancient hominin, post-ape minds? In essence, these brains would have primarily consisted of that integrated, multi-sensory, Dynamic-Core-based arena of experience (*aka*, the consciousness viewfinder), plus all of the mechanisms that we now consider to be the *unconscious mind.* At this point in mammalian brain development, creatures were already cross-coordinating & analyzing highly-sophisticated sensory input, using the cerebral cortex to store that incoming data in a large array of complex, modular & associative hippocampus-defined memories, and using that cortex to apply inborn & learned rules to emergent current & pinged-via-association stored data in order to help construct unique & dynamic behavioral responses—a cognitive loop that also produced neurotransmitter-fueled proto-emotions and was managed by organs like the thalamus & basal ganglia, *and* that used the amygdala to aid in emotional production/regulation & for ancillary storage of intense, primal pain- & fear-based memories. (This loop is depicted by our *Quick Sketch of Pre-Language Mammalian "Cognition"* in the Appendices.)

These are, indeed, the primary mechanisms lurking beneath & helping to sustain *modern* human consciousness. This old original hominin system is merely lacking the ability to "talk to itself" about what it's doing. Basically, the most elemental aspect of mammalian consciousness—the Dynamic-Core-based arena of experience—is a part of *all* vertebrate brains (going back to those *lampreys*). But until mammals, this conscious arena is only used to direct entirely reflexive & pre-programmed (aka, robotic) motor responses.

With the emergence of modularized experiential data patterns in mammalian brains, that conscious arena began to help direct dynamic & creative *learned* motor responses, allowing mammals to diversify their behavior based on their environment & experiences. And this kind of modular-but-non-linguistic mammalian cognition is exactly what advanced apes & early hominins employed. Humans have merely (& spectacularly) added a language-based inner voice that allows us to internally contemplate & manipulate that experiential data & those motor responses—aka, *self-awareness*.

I believe, in fact, that all of our mammalian (& probably all of our earlier vertebrate) ancestors did (& do) experience a type of non-language-based consciousness that is in many, many ways very similar to our own experience of consciousness. Yet, we are unable to truly fathom what the difference is between these experiences, because we simply cannot place our minds in a world in which there is *not* a word for *everything* (you can't even contemplate the *notion of it* without using words to contemplate it).

Ultimately, the complexity with which we are able to contemplate that huge amount of nuance, detail & interactivity in the world around us via our internal dialogue's system of language represents a *massive leap* in the nature of the conscious experience that is produced. And this perspective actually presents us with a pair of seemingly-contradictory truths. On the one hand, our language-based internal dialogue clearly sets the nature of our consciousness apart from the experience & capabilities of all other earthly animals. On the other hand, the essential mechanisms that define consciousness exist in a rather fluid continuum across vertebrate (& advanced *cephalopod*) species—and most animals do, indeed, experience a kind of consciousness that has much more *in common* with our own than it has differences.

How To Behave Like A Human

What does this really mean in terms of experience & behavior? It means if you kick a dog, the dog will feel (*experience*) pain, and sadness or anger, and it will know & remember that you (& the kick) are the source of its pain, sadness or anger, and it will both express this feeling in its behavior toward you and attempt to take action to protect itself from further pain—pain that its brain is cognitively *predicting* might happen based on both this & previously-stored experiences. That's some pretty heavy-duty lifting on the part of this little canine's primitive, limited consciousness.

And yet, the dog cannot *think* to itself or try to specifically *communicate* to you: "That makes me sad & angry because that both physically hurts & because I thought you liked me, and when you kick me it makes me feel like you don't like me, and I like you & I really want you to like me. And let me tell you why I like you..." The dog simply *behaves* in all of these ways (& experiences these feelings) based on those dynamic, modular cognitive systems that are making use of all those sensory input tools, stored data, motor scripts, switchboards, and behaviorally-governing, neurotransmitter-fueled emotional mechanics. (Of course, being humans, upon observing this dog behavior we are prone to naturally *interpret* it as the dog expressing that complex dialogue.) Those aforementioned mechanisms are all part of that essentially non-*thinking* system of pre-human consciousness that does everything but talk to itself.

This kind of brain & the resultant behavior are much like the "zombie" human that we contemplated in our first essay—a less creatively- & dynamically-efficient version of a human that does not possess a capacity for internal dialogue. Imagining that zombie helped us to demonstrate why it was evolutionarily advantageous for hominins to eventually morph this capacity for modular cognition into an ability to dynamically "tag" those modular elements with words & transform our behavioral rule system into a syntactic rule system—language—that could make use of those words to help generate much more creative & robust predictions & solutions.

When we look at the human mind this way—as a machine that is capable at its evolutionary core of operating in an almost fully-responsive capacity without employing internal dialogue—it is clear that the *vast majority* of our behaviors & actions can be controlled entirely via our *unconscious* mind. For example, think about reading an article and occasionally reaching for & drinking from a coffee mug while reading. In cases like this, your consciousness is almost fully-engaged in your primary task: the article's narrative. But when you glance at the mug or experience an internal pang of "coffee-desire" your attention might be diverted enough to momentarily think "coffee" between the other sentences in your head.

This data input (the sight or the pang) triggers both the thought "coffee" & related motor scripts that enact your physical coffee-drinking routine. This motor script is partly enacted because there is no conflicting unconscious inhibitory script also being triggered—like pain in your free hand that negates the act of reaching and might then force you to pause reading & consciously free your other hand from the magazine (or iPad) to reach for the mug.

If the coffee-drinking motor script is uninhibited & immediately enacted, your consciousness almost-simultaneously returns to the article. You're not thinking to yourself "grab the mug, bring it to my lips, sip"—at this point in your life, drinking coffee in this circumstance is a deeply-rote script and can be enacted without that kind of focused conscious attention. That’s because, according to our theory, once a deeply-rote, ongoing motor script like this is triggered, it can be maintained through a spatially- & physically-informed *unconscious* “action loop” that is routed to motor areas *without* engaging conscious cognitive processes. (This “action-maintenance loop” performs the same function in pre-language mammals, who can also use their more primitive cognitive processes to trigger dynamic motor responses that don’t require ongoing cognitive direction, but do require ongoing physical/spacial maintenance.)

Thus, as you drink the coffee—although your own unconscious act of drinking is immediately within your realm of "awareness," allowing you to peripherally perceive these coffee-consuming actions as you read—your own act of drinking may or may not become a *true* part of that conscious & remembered experience. This is why, immediately after finishing the coffee (which might've come as a complete surprise) if asked how many times you actually drank from the mug or at which points in the article you drank, you might have no idea—even though you were vaguely aware of every act of drinking that took place while you read.

And when we look closely at our lives & the actions that compose our days, it appears that most of it is actually a result of our unconscious mind humming away like it has in mammals for eons: responding to incoming data with a barrage of subconscious competing/cooperating motor scripts that have no need for internal dialogue in order to function and maintain sophisticated behaviors & actions. Thus, to actually behave like a human (as opposed to any other animal) is not really to *do* any of the things that we do—it's to do them, and *while* doing them, think something like: *"Man, that was stupid."*

Motor Task Chunking

Apologies for the *academic-paper-ish-ness* of our section heading here, but the poet in me (which is, oddly, how I began my brain expedition) could not resist the strange, thick music of the words: *Motor Task Chunking*. And the words do, indeed, describe exactly what we're going to discuss: the chunking of the consciously-attended-to "grab the mug, bring it to my lips, sip" into the almost-entirely-unconscious *coffee*.

The best place to begin exploring *Motor Task Chunking* (okay, I'll stop) is way back in those pre-language mammals (like dogs & monkeys) who use modularly-constructed proto-narratives in the process of recording high-priority experiences & building dynamic cognitive responses. In those early minds, in order for the modular components of this cognitive process to trigger actual actions (which is the whole point) those components would have to link to or trigger specific & appropriate motor scripts. (The thalamus' & basal ganglia's switchboards aid in transmitting & processing those often-competing cognitive & purely-reflexive motor scripts en route to the motor cortexes that coordinate their execution—something we'll discuss in more detail later.)

In humans, this cognitive process is overtaken (& powerfully, exponentially enhanced) by complex language. Nonetheless, the same fundamental relationship exists (as it *must*) between those narrative components & motor scripts. Thus, in humans, the words of our internal dialogue can (and often *do*) lead directly to actions. (Although—to give an awfully revealing *sneak preview* of our ultimate verdict on free will—the motor instructions that result from those words are likely triggered by the just-generated dialogue in the micro-moment *before* the dialogue's appearance within our Dynamic Core allows those words to be *heard* by us.)

In this system, the more complex & elaborate the motor script you can tie to a single word or thought, the more efficiently your consciousness can "off-load" the handling of full, multi-step motor sequences to those motor systems (sequences that our consciousness doesn't *really* need to be involved with). When we first learn a complex sequence, those full, multi-step motor scripts simply don't exist yet. Thus, we need to cognitively break that sequence into the smaller components for which motor scripts already exist.[5]

This means we have to actually *think* the words that trigger those smaller, already-learned components: *grab the mug*, *bring it to my lips*, *sip*. This is obviously a fairly inefficient way to drink something—and a pretty criminal misuse of those magnificently creative & robust systems of human consciousness. Our brain is *very* interested in getting a full, multi-step motor sequence in place for this mundane (but still very necessary) action so it can trigger it via that *blink-of-an-eye* "coffee"—and thus, keep its valuable conscious energy focused on more rewarding matters (like the *in*valuable insights coming forth from Sarah & Vinnie on the *97.3 FM Morning Show* that you listen to as you drink that coffee).

When we're *hoping* to trigger one of those fluid, multi-step motor scripts (usually one learned through intense practice) but *instead* end up *thinking* ourselves back into those less-fluid un-chunked component scripts, it has a specific & feared name: *choking*. This is why it's so bad when a Major League second baseman starts to freak out and *think* about the actual physical act of throwing the ball to first base. Suddenly that fluid, multi-step motor script is being interfered with by those cognitively-expressed smaller, component (& less fluid) motor scripts that are now actively *conflicting* with the more fluid one.

In an act as precise as throwing a baseball at the velocity that a Major Leaguer does, any tiny hitch or *oh-so-momentary* conflict in what muscle is doing what (because you suddenly can't stop *thinking* about the various aspects of throwing that baseball) can have professionally-disastrous results. This is also exactly why the *anxiety* produced by *worrying* about whether or not you can make the throw leads to the same problem that likely seeded the doubt (*all too often*: *one bad throw leads to another*).

Back in essay #2, we explained how those pain-based (or *survival*) emotions are designed to make our actions & thinking less fluid in favor of a more "hyper-aware" state. (Essentially, a decrease in focused, fluid attention is the cost of an overall lower-risk & diffusely-attentive state.) When we're nervous or anxiety-ridden, our brains are predicting a bad result from the upcoming action or event, and thus, want us to *slow down & think about this! And don't just stare at the problem—look around & see if you can come up with something better!* These are not the ideal neural conditions for throwing out the runner at first.

And this kind of problem is mostly a case of our *conscious mind* sticking its nose in where it doesn't belong—and where it *claimed* to supposedly have no interest, which is why we built the multi-step motor script in the first place. But our consciousness is like the ultimate *helicopter parent*—and as soon as it suspects that you're about to make a *valuable mistake* (the anxiety tattled on you), it has a tendency to step in and try to assert its ever-guiding influence over the whole matter. Although this can occasionally lead to some embarrassing scenes in front of your friends (or 40,000 baseball fans)—more often than not, when it really counts, the *hesitation* being counseled by our consciousness is *exactly* what the situation calls for.

When this whole process is working efficiently, however, it can allow someone like an experienced pitcher to calmly survey the batter, devoting all of his conscious mental capacities to the many nuances of this momentary conflict between them. Because he has learned & developed the highly-fluid & elaborate motor scripts necessary to enact widely-varied versions of throwing the ball, he can trigger a series of complicated actions via the tiniest part of a cognitive thought (& some closely-related, suddenly-applicable complex script might even step into the process almost *unconsciously*—like reflexively zipping the ball to first base when he ever-so-slightly detects the runner leaning a little too heavily toward second).

With his cognitive processes freely devoted to his engagement with the batter, the pitcher can bring all of that additional data to bear on the execution of the complicated, unconscious & fluid motor scripts. Instead of worrying about the actual act of pitching, he's calculating the nuances of the entire pitcher-batter conflict. It's the difference between a Little League pitcher & an experienced hurler. Without all of that practice in turning smaller component motor scripts into multi-step scripts—the complicated act of throwing different pitches to different batters in different game situations requires so much conscious cognitive work that there's simply *no room at the neural inn* for calculating anything like nuance.

So, besides how the words *sound*, that's what's so great (& occasionally vexing) about *Motor Task Chunking*.

The Diffuse Box Of Consciousness

In essence, this continually-dynamic conscious experience layered atop a deep, robust, ever-percolating well of unconscious activity is what Narrative

Complexity refers to as the *Diffuse Box of Consciousness* (a concept we've touched on briefly in the previous essays). I use the term "diffuse box" because although the experience of human consciousness can be specifically defined as one part of our internal dialogue loop, the content of that dialogue is continually & powerfully shaped by all of the data that is also being processed & managed within those same loops, but does not necessarily ever emerge in our internal dialogue.

Keep in mind (pun *sort-of-intended*) that the final singular narrative parcel we actually *hear* is the result of associating, comparing & culling *myriad* data patterns that have been pinged by the just-consumed linguistic, environmental & physical input. This means that even though you may not have *consciously noted* the painting of the sailboat at the dinner party (i.e., you were only peripherally processing that visual data while talking to the attention-consuming attractive attendee) you still *unintentionally* end up steering the conversation in a slightly different direction because you once almost *drowned* in a sailboat accident (due to how that peripheral-but-powerful data input impacts your unconscious data & dialogue *culling process*).

This might seem like a pretty sneaky & possibly harmful way to shape behavior, but it makes sense that our brains would seek to *slightly weight* our cognitive processes in one direction or another based on peripheral *currently-low-priority-but-historically-significant* data within our immediate environment. To make matters even more fluid, consider: even if the *first round* of processing didn't allow the data to escape the unconscious, the subsequently-pinged data processed within our loop can easily gain enough prestige or attention to allow that original unconscious thought's or action's *next generation* to emerge in soon-to-be internal dialogue. (i.e., Suddenly interrupting your conversation with that attractive attendee with an apparent non-sequitur like *"That sailboat painting is kind of freaking me out."*) Thus, consciousness is best described as something that both has distinct borders & a highly-fluid, interwoven-with-the-unconscious nature: a *diffuse box of consciousness.*

In a strange way, living within that box and inhabiting most of its *diffuseness* is the old hominin within us, dumb but not *dumb*—silent, but still capable of learning & enacting nearly *anything* that you teach or show or even ever-so-momentarily *request* of them. And floating atop the box—concretely & securely, but *barely-just-above* the diffuseness—is us, language-based human consciousness: a transformation of the most prestigious emergent diffuseness into a fluid, flexible, highly-integrated thread of thought & experience, a dynamically & complexly organized narrative equation born to be consumed again by the unconscious & help seed all the thoughts to come. We are the effervescent result of all the diffuseness percolating beneath us.

Imagine the loop of consciousness perpetually cycling, the majority of its data only either sparking other unconscious data via association or triggering primarily unconscious motor scripts. Based on the focus of our current actions, emotions, behavior, environment & attention, a tiny sliver of the highest-priority data running that loop ends up taking the exclusive & primarily single-narrative-only route into the arena of our conscious awareness, emerging in the form of *self-heard* internal dialogue—a thread of language-based thought woven into a complex Dynamic Core of experience. That route also includes a round-trip back to the unconscious processing in which it originated, where it rejoins the great unwashed masses of recycled thoughts & incoming data—some of which thanklessly results in those myriad unconscious actions, associations, emotions & behaviors.

From this perspective, we might think of the evolution toward modern human consciousness as a neural infrastructure initiative that slowly built both a new kind of high-tech engine to go around the same old track, and a new *extension* of the track that only accommodates this high-tech engine—a track that takes the engine to & from an exquisitely-evolved & high-tech consciousness-inducing destination within our prefrontal cortex.

Of course, the question we're really seeking to answer here: did that new engine bring with it *true* free will?

The Sliding Scale of Story, The Executive & The Virtuoso Switcher

Before we delve into the soul-defining answers to that question, let's explore a closely-related narrative mechanic (nostalgically taking one last brief *detour* in an essay-to-essay journey that has already taken generous liberties in indulging detours). This deeply-applied narrative mechanic might actually *only* be considered via detour—because its powerful influence on decision-making is hidden away in the very nature by which we tend to construct & consider the narratives that surround the decisions produced by our loop of consciousness.

This mechanic relates to how we instinctively (usually due to previous emotionally-biased—& thus, *recurring*—cognitive patterns or "habits of thought") frame the *scale* of these decision-generating narratives using our more advanced human predictive skills. We'll cut this detour right to the chase by providing a common example: the decision to photograph one's child as the child walks toward the school entrance on their first day of kindergarten. For our purposes, let's imagine this scenario confronting a naive first-timer dropping off their only child—thus, we will assume that this parent has no previously-stored (personal or observed) experiences upon which to base a more reflexive, habitual response to the moment. First, consider how an earlier version of the mammalian brain might approach this decision (forgetting for a moment that dogs & such don't even know *how* to take photos). That kind of dialogue-less system would be likely to frame this choice within a more narrow narrative scale that is ultimately unconcerned with the moments that extend beyond the very near future.

But humans—mostly due to those language-based, boundary-pushing predictive & future-imagining capabilities—are able to project from the current first-day-of-school moment to a future moment in which they will want to remember & re-experience (& *not lose forever*) the *emotion* of this moment. According to our theory, this *self-projection* into the future is ultimately another one of those predictive cognitive tasks, and thus relies on those linguistically-based tools of internal dialogue. Because our first-time parent has no previous similar (thus likely to be pinged) experiences, they will need to consciously consider that they might want a photo later in order to arrive at the choice to take the photo now. Since it is not a rote & automatically triggered motor script under these conditions, the act *cannot* be triggered without generating the desire & possibility through internal dialogue contemplation. This act is a unique, new choice in an unfamiliar environment & circumstance, and thus is likely to rely heavily on the creative-choice-generating capacities that our internal dialogue mechanisms specialize in. Without contemplating & articulating (internally or verbally) the future desire for the photo, there is no reason in the current moment to actually take the photo—since (if we imagine our scenario in that *real-rolls-of-film* pre-Facebook era) it provides no in-the-moment pleasure nor satisfies any other in-the-moment need. This is why almost no non-human brain would even *consider* taking the photo—it can't *imagine* any reason to.

Within most non-human brains (except probably mammals like dolphins, elephants & advanced apes) that *in-mind* future version of the animal—and most of the imaginable possible moments that might occur between now & then—don't *really* exist to them or are even *available* to their minds for genuine creative & manipulative contemplation. Those other mammalian brains *do* use prediction patterns to help calculate & select most-beneficial responses based on considering predicted near-future results. Nonetheless—even if the predicted result requires multiple steps to achieve—this is really an extrapolation of symbolic & associative causal logic, not a true creatively-conjured & *mentally-malleable* internal self-depiction of the future.

By using this kind of sliding scale of story, our brain essentially views every decision as a "hinge" along a narrative continuum—preceding the hinge is the narrative history leading to this story moment, and following the hinge is the predicted future of this narrative. In making a decision, the brain seeks to inhibit or activate this hinge in a way that is most likely to result in a desired future narrative (or to avoid an undesired one). Thus, how we internally conjure, perceive & scale both past *and* future elements of the current narrative impact how the brain chooses to flip the decision-hinge.

This makes the scaling of story an *essential part* of most decisions that we consciously contemplate via internal dialogue. For example, if a cigarette smoker views the act of smoking within the smaller-scale narrative of the stress of the previous/current moments and the pleasure they currently desire/will soon receive, the smoker will very likely choose to light the cigarette. But if the smoker in that moment expands—via internal dialogue—the scale of that narrative to include the likely future prediction of themselves dying young from lung cancer, they might be more hesitant to light the cigarette in that instance (or be quick to add a new twist to their predicted future narrative, like imagining *quitting next week* in order to inhibit the inhibition to smoke produced by the thought of dying young). Of course, in the latter case, even a smoker whose brain doesn't bother to add a new future-twist might very well light the cigarette without hesitation, which is part of the problem with smoking & its sinful kin: they produce some *very* powerful desires for their temporary pleasures. Which conveniently brings us to our next stop along this detour: the battle between urges & narratively-reasoned desires.

Back in essay #2 we noted that those urges (mostly based on pure physical desires or threat responses—aka, our human versions of those ancient "proto-emotions") tend, at their very highest levels, to outweigh most opposing narratively-generated desires in our decision-making process. Whenever we are starving/parched, deep in the throes of lust, completely exhausted, totally repulsed by disgustingness, or in the grip of fight/flight (*or seriously craving a smoke…or really, really angry*) the desire to satisfy that powerful urge will often influence our action-choice more strongly than any narratively-generated (and/or belief-based) impulse to act otherwise.

The ultimate results of these kinds of consciously-contemplated decision-making battles appear to be primarily determined by our “executive control” area, the *dorsolateral prefrontal cortex*[6]. Other areas might be involved in helping the DLPFC to mediate this process, but it’s generally viewed as the decision-making hub. (*And these battles can also be impacted by that endorphin-based willpower discussed in essay* #2). Although an urge (or a *craving*) can originate via unconscious sources, once we become consciously aware of the urge, it can be contemplated via internal dialogue —allowing our choice to *inhibit or encourage* the urge to be handled by that executive control area (which our narrative building/analyzing mechanisms feed data into). And in the view of Narrative Complexity, our language-based parcels of internal dialogue (built by our left hemisphere narrative-building mechanisms) must be "processed" in two key ways *before* the DLPFC is able to use those narrative parcels when deciding to inhibit or encourage an urge or action.

One: the dialogue needs to generate—likely via categorical, associative (& habit-driven) methods—a specifically-selected (potential) action trigger (or a set of *cooperative triggers*) that might aid in satisfying our urge or narratively-based desire. Thus, this associatively-based *potential-action selection process* can use an (anger-induced) thought like "I'm gonna kill him!" to help select a goal-specific (& *emotion-satisfying*) potential-action trigger among multiple options—like reaching out to strangle the target or seeking out the nearest heavy object.

The anger (& the supporting threat-response proto-emotion *fight/flight*) that *induces* & helps to define the *intent* of this potential-action-triggering thought is *initially* produced by the previous (& likely observational) thought—something like, "He's trying to hurt my child!" (*therefore*, "I'm gonna kill him!"). And if we really want to achieve our *malevolent intent*, our chosen action trigger had better be capable of cuing a well-practiced, elaborately-chunked, complex & highly-fluid motor script.

In addition to spurring this *potential-action selection process*, every just-built parcel of dialogue must also be analyzed according to all of those myriad *emotional & proto-emotional equations* (which judge our 13 primary emotional pairs & various physical urges). As explained in Essay #2, these "equations" make a wide range of judgements about factors like value gain/loss, prediction success/failure & belief compliance/violation (all of which are involved when analyzing a narrative parcel like "I'm gonna kill him!"). Consequently, this process of emotional & urge analysis involves a wide range of brain areas, such as (but not limited to) the orbitofrontal cortex, the anterior cingulate cortex, the amygdala, and the insula—all of which could process this narrative/emotional data in parallel (although it wouldn't be surprising to find interaction between *some* of these areas during processing). And this emotional analysis occurs at the same time as that *potential-action selection*.

Therefore, we hypothesize that the DLPFC receives our language-based narrative parcel (with its newly-attached & just-selected potential-action trigger or triggers) in conjunction with this wide variety of inhibiting & encouraging emotional judgments produced by that wide array of emotion-analyzing brain areas. The executively-controlling DLPFC is then able to compare & calculate separate-but-relevant emotional judgements—such as predicted value gain/loss (excitement/fear), prediction success/failure (confidence/anxiety) and belief compliance/violation (satisfaction/guilt)—allowing all of the varied emotional factors to contribute to our decision to encourage or inhibit any narrative parcel's potential-action trigger(s).

Our model of emotion's role in decision-making is similar to the model that Oxford neuroscientist Edmund Rolls proposes in his 2014 book, *Emotions and Decision-Making Explained* [7]. One key difference, however, is in the placement of action selection processes. Although Rolls' model places action selection *after* emotional analysis, in our view, all of the necessary data required to select an appropriate action trigger (i.e., intent, action goals, action targets & action obstacles) is present in the syntactic & vocabulary content. Thus, because this is the same data that's used for emotional analysis—and because emotional judgements do not seem necessary in *selecting* a potential action—it seems more likely & efficient for action selection & emotional analysis to occur simultaneously.

In addition, from our perspective, if the emotional judgements that ultimately determine action activation/inhibition also contribute to the action-selection process, then our brain would already have all the data necessary to decide activation/inhibition at the time of choosing the action. This means there wouldn't technically be any need to ever emotionally *inhibit* any potential action choice, because we *already know* what we're emotionally willing or not willing to do when we choose our action. In such a scenario, why would our brains ever choose an action that it *already knows* it isn't emotionally willing to enact? This would also seem to make the DLPFC less of an executive and more of an *executive assistant*.

In our model, the same raw data—syntactic & vocabulary content—is processed by our emotional analysis & action selection mechanisms simultaneously, and the results of both are reunited in the DLPFC to determine whether the action is activated or inhibited. Basically, if the emotional judgements that support or encourage any urge- or narratively-based possible-action(s) neurally *outweigh* competing emotional judgements that discourage or inhibit the potential-action(s), then the selected action trigger(s) are "chosen" to be enacted (or vice-versa, preventing or inhibiting the selected action). Ultimately, from our theory's perspective, this decision-making process is not usually a traditionally-conjured "this-action-or-that-action" choice. Rather, it is a choice between enacting or inhibiting a singularly-directed action (or set of actions) that satisfy a specific urge or narrative desire.

In other words—returning to our homicidal example—when someone else's anger-inducing act sparks the impulsive thought "I'm gonna kill him," we aren't choosing between strangling him or, say, merely spitting on him. The language-based narrative intent spurs the selection of an appropriate urge/narratively-satisfying action trigger, e.g., reaching out to strangle him. The DLPFC's job is to then use the proper emotional judgements to choose to enact or inhibit that singularly-directed action (or set of actions). Thus, we are choosing to strangle him or not to strangle him in that moment, not choosing between strangling him or taking some other specific-but-non-lethal action. Of course, because our consciousness is a looping *speed demon*, we can inhibit our desire to strangle & follow it up with an uninhibited desire to *expectorate* so quickly that it basically & perceivably *feels* like we're choosing between the two (unpleasant) options in the same moment.

~

When examining how we choose between activating or inhibiting all of these varied *desires*—vengeful, romantic, culinary, material, *ad infinatum*—there's another key question that needs to be asked: how exactly does our brain define *how much* we desire any particular thing or result (or *person*) in any particular moment? We hypothesize that there are two primary factors that combine to define our level of "desire" (aka, urge/emotional *intensity*) for any particular thing or result. The first factor is the syntactically- & vocabulary-defined overall *value* of the thing or result—determined via those three "Narrative Prioritizor Test" criteria we identified in Essay #2: importance, relevance & novelty.

Importance essentially represents that thing/result/person's experience-based *personal value* to you. In other words, if you've learned to love comic books, your brain might judge comic books as having high personal value—which increases your *desire* for lots of those particular kinds of things/results/people. Because they are based on our experiences, these importance judgements are highly individual, and *anything* can become highly important to *anyone* if their life experience has made it so.

Relevance represents the thing/result/person's "closeness to you" in terms of various in-group/out-group status judgements. If your goal impacts *any* of the social groups that you consider yourself a part of (family, school community, neighborhood, *ad infinatum*) then it has *some* level of relevance to you. So if your comic-loving self is also a member of a super-tight comic book club, or if comics are a cherished tradition that you share with your dad, then you might judge any comics-related goals/tasks that specifically impact them as *highly-relevant*. The greater the role of that social group in your life, the greater the relevance, and thus, the greater your *desire* for achieving any particular goal that benefits or relates to that social group.

Novelty represents the judgement: how many of these particular things/results/people do we *perceive to be* available vs. other things/results/people in this same category. For example, how many of these particular comic books (*Spider Man #1*) do we perceive to be available vs. other comic books in this category (*all Spider Man comic books*). The higher the novelty judgement, the greater our *desire* for that particular thing/result/person. Here again, because our *objects of desire* are heavily shaped by personal experience, the way in which we categorize those objects and how much we value their novelty is highly individual.

As noted, all of these value-defining criteria can be employed by applying a narrative parcel's syntactic & semantic content to our emotional equations. This "overall value" definition of things/results/people is the first factor in determining *how much* we desire anything in particular. The higher the overall value, the greater the desire. Thus, the comic-lover from our previous paragraphs might most-greatly desire & cherish a rare comic book received from (or possessed by) their collecting-partner dad.

The second primary factor in determining desire is something that we refer to as "emotional attenuation"—which is a cousin to what Rolls' dubs "sensory specific satiation." As Rolls describes, in the process of satisfying a general urge like hunger, we experience a satiation of sensory-specific stimuli (like a particular kind of taste) throughout the process—which drives us to diversify what we consume as we satisfy that hunger (helping us to meet different dietary needs). According to our theory, a parallel "satiation" or "attenuation" mechanism (a decrease in stimuli response intensity after increased exposure to that stimuli) is also present within our *value-based* emotions. (Such attenuation is *not* likely present within our belief-based or purely validity-based emotions—because greater exposure to belief- & pattern-prediction-supporting stimuli typically *strengthens* our subsequent response intensity, i.e., *increases* the likelihood of belief compliance, or our degree of confidence in that pattern prediction.)

There appear to be two kinds emotional attenuation mechanisms: general emotional attenuation (e.g., attenuation of general affection) and category-specific emotional attenuation (e.g., attenuation of a specific source of affection). This is, essentially, how our brain judges our "appetite" for (or our *level of deprivation* regarding) any particular thing/result/person in any particular moment.

For example, if someone has an active social life with lots of friends & family and is a frequent recipient of their abundant affection (aka, lots of exposure to that "general" emotional stimuli) the prospect of a mid-week social gathering might not generate an urgent desire to attend (aka, a decrease in response intensity). Conversely, if someone who's starving for affection (minimal exposure to that general emotional stimuli) receives the same invite, they're more likely to have a stronger desire to attend (a more intense response to that general emotional stimuli). These are both demonstrations of general emotional attenuation mechanics at work.

Now imagine that the affection-abundant person is provided that same invite, but in this case the gathering will an include an old best friend that they haven't seen in years. Because this person's affection-response has been attenuated (by high levels of general affection) the emotional affection stimuli must be fairly significant in order to generate an intense response (aka, a stronger desire to attend the gathering). Luckily, the presence of the long-lost best friend provides category-specific affection (the affection of a dear old friend) that this person *has not* been heavily exposed to. Thus, this category-specific emotional stimuli is still capable of providing a more intense response from our affection mechanisms—leading to a greater desire to seek out that affection at the suddenly-more-exciting mid-week gathering. However, even in a case like this one, if our "general affection appetite" has been *super-satiated* we still might not be able to generate enough desire to motivate our attendance (*I'd really love to see Bob, but I'm just too burnt out from all the socializing— I really have no desire to go*).

Conversely, and sadly, if someone is deeply affection-starved, even an ultra-redundant category-specific low-value emotional stimulus—like a thrice-daily 10-second social interaction with a doorman whom you know is only really pretending to like you—can still generate a more intense emotional response, and thus, a stronger desire to ensure that the interaction always occurs thrice daily, *ad infinatum*.

The mechanics of attenuation, desire & decision-making also have significant implications in regard to both behavioral *and* chemical addictions—which are clearly major problems in modern society. In fact, addiction issues beg for an entire essay focused on them. But *Free Will* still beckons, so instead here's a quick glimpse into the unique difficulties posed by *chemical* addictions… Imagine that your stomach actually grew a little *every time* you got carried away with the joy of consumption and over-ate. It's easy to see how an infinitely-sizable stomach and its subsequently infinitely-sizable appetite could result in us consuming far more food than the rest of our systems could adequately process. This is the problem that chemical addictions produce.

Because the attenuation of a purely chemical appetite (like opiates) has no physically-correlating boundaries (like a stomach), and because the amount of the chemical available has no production limitations (like those that physically limit the production capacity of internally-generated chemicals like endorphins)—our external chemical appetite and the amount of chemical available are nearly infinite. I say *nearly* infinite because there is one very firm boundary to our chemical appetites: death. In the end, sure, emotions like affection are powerful stuff & we can get a great buzz from large doses—but we're never going to love something so much that we generate a big enough buzz to accidentally kill ourselves. The same cannot be said about those chemicals that so many of us have grown to love so dangerously much.

One other tragic attenuation-based side-effect that relates to opiate addiction in particular: in the view of our theory, the addiction itself drastically lessens our internal capacity to exercise neurally-based *willpower* when attempting to choose *not* to satisfy that infinitely-growing appetite. As discussed in Essay #2, we theorize that our willpower capacity (which is a key factor in our ability to use narrative motivations to resist an urge) is heavily dependent on endorphins—and endorphins make use of the same kinds of opiate-based receptors that the opiate-based drugs are severely *over-attenuating*.

In other words, the same neural mechanic (attenuation) that's causing the addict to require larger & larger doses in order to achieve a response (*get high*) is also causing the brain to require larger & larger quantities of endorphins in order to engage our willpower. The major problem being that (as we noted) our brain actually has internal physical production limitations for chemicals like endorphins—so once that opiate-appetite has outsized our internal capacity for endorphin production, our willpower is utterly toothless.

In a scenario such as this, an individual's willpower capacity is essentially reduced to zero. These are likely some of the reasons why opiate addiction is so deadly, and why the widespread availability of powerful opiates in our society is pure madness. (Infuriatingly, pharmaceutical corporations' desires for increasing profits far exceeds their willingness to behave responsibly in pursuit of those profits. In particular, the Sackler family and their company Purdue Pharma have been historically vile in their greed-driven disregard for how opiates can destroy lives.) This is also why it can be crucial to provide access to "weaning" drugs like methadone—because minimally-satisfying the drug desire with smaller dosages of less-potent chemicals allows the brain to slowly *lower* that level of opiate attenuation (without the interference of intense cravings—*aka, withdrawal*). From our view, this helps to provide those endorphins with a fighting chance to effectively engage those willpower mechanisms in the battle against the self-destructive behavior.

Returning to our exploration of a *neurally-healthy* decision-making process, fortunately, most of our emotionally-motivated choices and behaviors are not nearly as prone to total willpower failures—although they *are* prone to smaller ones. In fact, humans are notoriously fickle creatures whose willpower to resist or succumb (aka, *impulse control*) frequently flickers between inhibition & activation on an ongoing basis. Thus, even when a narratively-expressed intention & its selected (potential) action are initially inhibited, they can still have a powerful influence on our *subsequent* action-decisions. In other words, going back to that potential strangler, even if they inhibited their *initial* impulse to reach toward their target (which was generated via their anger-induced response "I'm gonna kill him")—they still might very well a *want* to kill their target.

This means that if, for example, the primary emotional and/or belief judgement(s) that initially inhibited that action weren't very strong (e.g., a belief that *Murder is only bad sometimes*, instead of *Murder is always bad*) then the action (or a closely-related action) might have a good chance of being *uninhibited* in a subsequent round of processing. In such a case, the subsequent parcel of dialogue that generates & accompanies the action might only be tangential (linguistically) to the original intent, and not directly intentional. For example, our aspiring strangler might simply think "Screw it!" as they now activate their just-inhibited strangling motor scripts. (Instead of thinking a more directly intentional "No, really, I'm gonna kill him.")

In our brains, "Screw it!" can also mean "I'm gonna do this!" (and carry that specific action-triggering intent) because in our narrative construction the brain has actually selected "Screw it!" *specifically to represent* the intent "I'm gonna do this!" That's because—as explained back in our discussion of *narratively contextual rule application*—the *intent* of an under-construction narrative parcel is typically shaped & directed in part by the syntactic content of the *previous* narrative parcel (when it's not an entirely new narrative thread).

Thus—even though that first impulsive urge to strangle was suppressed— *in response* to that just-experienced thought "I'm gonna kill him!" and its intent, the continuing anger (& a focus on its target) can help to direct the generation of another (& possibly more intense) unconscious impulse to overcome the inhibition and try strangling him. The brain then linguistically expresses this desire (which it has identified as a pattern in the unconscious right brain emergent data and is processing under conditions & intent partly defined by the previous narrative parcel) as "Screw it!"—because it has learned to symbolically & habitually express such an impulse in that linguistic fashion.

And if the action was initially inhibited by something flimsy (like a weak belief) then something as simple as a small increase in anger when expressing "Screw it!" could help to tip the scales in favor of enacting the previously-inhibited but now "re-triggered" potential action. Before you know it, our momentarily well-behaved *potential* strangler is suddenly (& prosecutably) an *actual* strangler.

Our brains can also tip the scales in favor of eventual action (after an initially inhibited impulse) through the *internal* use of that selected-but-suppressed motor script. As demonstrated by examples like free throw shooters who perform better after mental rehearsal, triggering a motor script while suppressing physical action allows us to internally practice—essentially, imagine—the selected action. In emotional (& thus, decision-making) terms, this practice can provide us with additional *confidence* in taking successful action (and allow us to further calibrate those motor scripts based on imagined predicted results). What are we doing when we pause before jumping across a puddle, swinging our arms, bouncing our knees & picturing our leap? What's a dog doing when it whines below a higher-up landing, springing slightly with a few mini-hops before launching itself upward? We're both building our confidence in taking successful action—internally practicing (via our premotor & somatosensory cortexes) motor scripts that have been selected, but suppressed due to judgements like anxiety or fear (or in some cases, simple otherwise-motivated intent).

Together, all of this speaks to the surprising power (& usefulness) of selected-but-suppressed actions—which can serve both as a silent instigator who recruits emotions like increasing anger to motivate subsequent action, and as a quiet coach who helps us run enough practice drills to provide the confidence to go ahead and act. (Selected-but-suppressed actions can be *so* powerful that we built our entire loop of language-based human consciousness around a selected-but-suppressed action: internally-expressed, but unspoken, speech—aka, *internal dialogue*.) The power of selected-but-suppressed motor scripts is also why—if you're really hoping *not* to lash out with an uninhibited angry action—it's probably better to focus on doing *something else* instead of focusing on *not doing* that ill-advised thing. When we focus on specifically resisting an action (by repeating a thought like "*don't lash out*" or "*don't look down*") we're actually helping that selected-but-suppressed action to stay within our *potential-action* cue—thus providing the inhibited action with more opportunities to subsequently become uninhibited.

Rather than forcing our brain's *executive* to continue inhibiting a powerful urge every time you think "*don't lash out*" or "*don't look down*," it's likely more effective to focus our mind on sustaining a preferred *alternative* action. For example, avoiding an angry action by repeating a simple alternative action-oriented thought like "*just stand here*" or better yet "*just take some deep breaths*" (which is better because physically slowing and calming ourselves can impact how we *feel*—possibly helping to dampen the strength of emotions like that *scale-tipping* anger).

In a system such as this, our DLPFC's *executive* is less of an arbiter who considers various arguments, and more like a carnival strong-man hoping to slam his hammer with enough gusto to ring the bell at the top of the pole. Action-encouraging emotions like confidence, excitement, pride (and *anger & flight*, which cleverly employ endorphins to generate action from their inhibitory pain)—these help the strong man hammer with gusto, and action-inhibiting emotions like anxiety, fear & guilt can weaken his swing. In other words, those action-encouraging emotions can help to open & excite the neural pathways that carry action triggers forward toward our motor cortexes, and negative emotions can inhibit & dampen those neural pathways. (Rolls describes this kind of decision-making mechanism as "a single winning neuronal population.")

Although our wide array of variously-contributing (& often *opposing*) emotional judgements can make this decision-making process exceedingly-complex in humans, the process has roots in (& still resembles) some very ancient & rudimentary behavioral mechanisms. In both roundworms & humans this "action-triggering" mechanism does the same thing: it provides the capacity to override/interrupt current behaviors or actions and generate different situationally-appropriate behavior or actions when a new possible value gain/loss (i.e., a yummy or a threat) is identified.

Roundworms squiggle along according to rote, repetitive motor scripts until sensory organs identify bacteria and help generate serotonin- & dopamine-based neural responses that trigger ingestion motor-scripts, which override/interrupt the suddenly-inhibited locomotion motor scripts. Humans dance feverishly according to unconscious, musically-calibrated motor scripts until their eyes glean the delicious & thirst-quenching cocktail, and the sight helps generate urge- & narratively-based neural responses that trigger walk-to-the-table motor scripts, which override/interrupt the suddenly-inhibited dancing motor scripts. In humans, if an action trigger receives enough emotional support (or fails to generate *any* emotional discouragement) those motor instructions are carried forward via those switchboard mechanisms within our thalamus & basal ganglia (which possess connections to *task-selecting* cortical areas like that prefrontal cortex [8,9]).

All of our competing & cooperating motor script triggers—sparked by both the unconscious & internal dialogue narratives—end up sending their variety of competing/cooperating instructions through those switchboards. The switchboards then do what they've done since the beginning of vertebrates: manage & mix all those inhibit & activate messages so that the most-currently-appropriate result is eventually generated within the body's nerves & muscle fibers (via the thalamus & basal ganglia's connections to areas like the premotor & primary motor cortexes [10,11]). And tucked within this whole process there can be an additional & fundamental action-trigger that accompanies a narrative parcel: the act of actually speaking those words aloud, instead of just hearing them in your head.

It may appear, based on the description above, that those switchboard mechanisms are actually where the final action *choices* are being made—since this is likely where all of those *varied-source* motor triggers are either *shunted-away-from* or *allowed-to-travel-to* our motor cortexes. To help clarify our theory's view of the process, I'll provide a (tragically over-simplified, but still useful) analogy of our thalamus & basal ganglia's switchboard mechanisms…

Imagine a virtuoso "train-track-switcher"—someone in charge of a massive train station with myriad tracks traveling in & out. Some of these trains carry action-triggering instructions, and each of those trains needs to have its incoming track switched over to the proper outgoing motor-area-destined track (*if available*). Most importantly to our virtuoso switcher, each incoming action-train arrives with a specific "prestige" (or *priority*) value—and the switcher's ultimate goal is to make sure that the highest-prestige action-trains get first & best access to their necessary outgoing tracks.

These trains originally depart from myriad brain areas—like our sensory cortexes, amygdala, and that decision-making *prefrontal cortex*. And although, in most typical cases, our highest-level urges can override our narrative desires—if you possess the right motivation & enough willpower, that narrative desire can still win the decision-making battle. Thus, when a high-priority action-train arrives from those PFC-based *executive control* brain areas, it ultimately seems to be capable of superseding the prestige of *any* other competing action source—as demonstrated by those *most-excruciating* fear-&-pain-overcoming and executively-ordered actions, like *cutting off your own arm to save your life* (which requires those *executive* instructions to supersede the amygdala's *reflexive* protective response).

This is how an executively-directed (*consciously contemplated*) choice can be essentially inevitable *before* it reaches our switchboards, because its instructions are typically given highest-prestige treatment within those switchboards. Any executively-directed instructions can, however, be nearly instantaneously *interrupted* & replaced by one of those reflexive protective responses if, for example, our survival is suddenly threatened by a rapidly-approaching object (an object that was not contemplated in the momentarily-previous & conscious action-decision).

There are also plenty of lower-prestige instructions that are constantly allowed to pass through our switchboards & be enacted by our motor systems because they simply don't conflict with any of the "outgoing tracks" that are required by any currently executively-controlled actions (e.g., unconsciously walking & drinking your coffee while consciously thumbing a message on your phone). In all of these different cases, our switchboards aren't making the kinds of analytical, consciously-directed action selections that we associate with real *choices* or *decisions*—these switchboards are simply (& complicatedly) routing *previously-determined* instructions according to their *already-assigned/calculated* prestige/value and current motor-resource availability. (And these switchboards are also likely helping to route those internally-rehearsed *selected-but-suppressed* motor instructions to destinations like our premotor & somatosensory cortexes.)

In the case of lower-prestige, unconsciously-maintained actions (like drinking the coffee) it's also important to note that the data helping to direct or guide those actions is much more *generalized* than the data that helps guide consciously-maintained actions. And this *data generalization* leads to those many slightly strange or misdirected action results that litter our days. For example: You're sitting in the living room reading the paper on your iPad when you suddenly realize that you forgot to give your children their vitamins before school. *I'd better take those out & put them on the kitchen counter so I remember later.* This consciously-directed decision leads you to stand & start walking toward the kitchen, iPad still in hand…

Once you start walking and re-engage consciously with the iPad (because *of course* you keep your head buried in your tablet as you walk) those consciously-sparked vitamin-directed action tasks essentially leave our conscious loop and slip into that "action-maintenance loop" we identified earlier. And in this loop, the actions are no longer being directed by that more specific consciously-contemplated end goal (*put the vitamins on the counter*), but rather, by more generalized & immediate action-defined goals that the conscious end goal first sparked (*walk into the kitchen & open the cabinet*). Thus, these now unconscious actions are guided by that more generalized environmental/physical data regarding a path to the kitchen & a target for opening (research suggests that these kinds of "sensorimotor intentions" are specifically handled by the *posterior parietal cortex* [12]).

This is why, once you're in the kitchen and lift your head from the iPad to scan the open cabinet, you might be baffled momentarily (*what am I looking for?*) before remembering the vitamins and suddenly realizing *you're not even looking in the right cabinet*. When you first stood up from the couch, the pre-conjured task of opening the cabinet might've led you to actually open your free hand slightly in preparation for grasping the cabinet handle—leaving those unconsciously-looping instructions on continual hold until the right environmental/physical data triggers the rest of the *well-chunked* cabinet-opening motor tasks. But in our action-maintenance loop, this very *generalized* task-triggering data (e.g., the general physical attributes of a cabinet handle) is *no longer connected to* (nor does it *contain*) that more specific end-goal data (vitamins). Thus—even though you have absolutely no doubt about which cabinet the vitamins are in—because your action-maintenance loop only needed *cabinet handle* environmental/physical data in order to let loose the rest of its *open-the-cabinet* motor scripts, you simply (& stupidly) opened the first & closest cabinet noted within your visual arena.

And—partly depending on factors like whether you cued-up that grasp & open task by opening your free hand when you stood—you might get to "*what am I looking for?*" as soon as you find yourself standing dumbfounded in the middle of the kitchen. In this case, *opening the cabinet* never even made it into your action-maintenance loop (or it wasn't part of an elaborately-chunked motor task)—leaving that loop with no further unconscious instructions to let loose once you've reached the end of your walking path. This is another example of how our mind can unconsciously maintain & direct motor actions via more generalized information *without* having any conscious awareness of how (or *if*) that action is related to our more specific (& originally-consciously-defined) end goal.

These seamlessly (although *sometimes misdirectedly*) interweaving motor instructions occurring in response to widely-varied types of data from widely-varying sources are the ultra-complex descendants of those simplistic, ultra-ancient inhibit & activate behavioral mechanisms (observed in our dapper roundworm, C. Elegans) now grown into Godzilla over the course of evolution. And another point that this all reinforces: although we have been primarily talking about the central loop of consciousness that produces internal dialogue, as a whole the many processes we've discussed in these essays involve multiple loops & offshoots (that we *haven't* discussed) that merge & diverge from that central loop. For example, after that neural moment of potential action selection, that data might also be sent to motor cortexes—priming them for potential action execution at the same time the DLPFC is determining action activation/inhibition. There is *a whole lot* happening at each step through this neural maze—we're just traversing the main perception-to-action thoroughfare. [13]

To continue bringing our essay-to-essay journey full circle, this internal neural battle between competing & cooperating instructions is reflected in a description we offered in our first essay: *Deep down in our psyche, these are the kinds of impulses that are battling for our brain's undivided attention. Each moment of existence is a Roman Colosseum in our minds—each urge, each impulse, each desire tossed into the arena, fighting viciously to be heard, to be made part of the story, to be expressed out there, where the thing that thinks them acts its act in the world.*

Revisiting The Great Syntactic Divide

Admittedly—if your hope is that yes, humans *do* possess free will—the evidence presented so far by the *scientific* approach to this question is not encouraging. In terms of that chicken & egg conundrum, science is clearly leaning toward the side of the egg: our brain seems to have the capacity to *decide* to act (choose to activate/inhibit an action or task) slightly *before* that action's accompanying thought actually emerges in our consciousness. Returning to our act of drinking the coffee while reading, it seems clear that the coffee-sight or coffee-pang input could at least *simultaneously* spur the act of reaching *and* the internally-experienced thought "coffee." So by the time we *hear* the thought "coffee," the instruction to reach has likely already been given (or the task has at least been *chosen* to be subsequently—and essentially inevitably—enacted).

Consider that in terms of our loop, that parcel of internal dialogue's (auditory-cortex-aided) emergence within our Dynamic Core is that dialogue's *last* destination before returning to our unconscious processing. Thus, this neural moment must occur a few steps *after* that dialogue is processed by those (left-hemisphere-originating) narrative-building & analyzing mechanisms — which are necessary to first *create* & *evaluate* (syntactically & emotionally) the dialogue that we ultimately hear.

As we described, those build & analyze mechanisms are nearly *immediate* predecessors to neurally *enacting/inhibiting* that dialogue's accompanying (*possible*) action triggers via our DLPFC—because as soon as that action-triggering narrative parcel has been emotionally analyzed (for gain/loss, beliefs, etc.) the DLPFC has all of the data it needs in order to executively *enact or inhibit* those potential action triggers. Therefore, in a system such as this, our executive neural processes could receive all the necessary decision-making data *prior* to any accompanying dialogue's *actual appearance* within our Dynamic Core.

Nonetheless, there still might appear to be a tiny opportunity for free will to make a comeback in our argument. This is because thus far we've mostly explored more impulsive acts like lashing out angrily or reaching for the coffee, *not* more deliberate acts like reading the article—*aka*, circumstances where our consciousness is more focused, and therefore where it is most likely to exercise any capacity for true free will. Exploiting this tiny (& seemingly final) opportunity for free will's existence means asking the question: does the *conscious contemplation* of a task-triggering choice (via action-specific internal dialogue in a decision-making context) somehow make us *consciously aware of* that choice *before* the action is executively (& inevitably) chosen to be enacted? (Thus giving us some possible *agency* over that choice—*aka*, free will.)

This hopeful space within the loop where free will might yet exist is tucked inside the same (left-hemisphere-originating) narrative-building/analyzing mechanisms that we've been discussing. It's the location that we identified in essay #4: *The Great*

Syntactic Divide. Just in case (as it does for me) this term still only brings to mind John Belushi's delightful, garish mug (via the 1981 film "Continental Divide"—yes, large portions of my right-hemisphere are filled with peculiar data-associations) I'll briefly re-locate this epic juncture in our internal dialogue loop. The Great Syntactic Divide is that left-hemisphere neural moment *after* your parcel of narratively-based dialogue is actually constructed (via rule & vocabulary application to the highest priority pattern of emergent right-hemisphere data).

As described, this adjacent post-construction moment is when all of those emotional equations—including emotion-producing beliefs— are applied to the just-built narrative (a process that ultimately involves circuits with myriad brain areas) in order to help generate & send-off the proper instructions to the widely-varied emotion-producing portions of the brain, *and* in order to send those results to the DLPFC for use in decision-making.

We've hypothesized that these emotional equations *must* be applied here because—based on neural judgement principles supported by Daniel Kahneman's *Prospect Theory* equations[14,15]—we know that the brain's emotion-producing calculations are based on those contextual/narratively-defined relationships between value & validity. Therefore, the brain *can't* apply its emotional equations until *after* this data has been constructed as a narrative. Beliefs are also specifically-designed to make judgements about narratively-constructed data, and would therefore naturally need to be applied *after* this Great Syntactic Divide as well.

Where Are You, Free Will?

What opportunity does this provide for true free will's existence? Not much. If we genuinely have *agency* over our choices (dialogue-based *conscious awareness* of choosing to act prior to the executive—and inevitable—neural *activation* of that action) then internal dialogue would need to *circumvent* this Great Syntactic Divide, sending special & specific *action-impacting* dialogue immediately into our Dynamic Core to be heard *before* the brain undertakes all of that decision-determining emotional analysis. This would seem to be the only way to become *aware* of an inevitable choice to act *before* our executive machinery receives the emotional data that determines whether any action is activated (or set into *inevitable* motion). And this does not seem to be a very plausible data pathway for our internal dialogue.

In fact, our theory has already provided some very strong evidence that internal dialogue *must* be emotionally analyzed before it is actually heard by us or spoken aloud. This is because all dialogue—even sentences that we only hear in our heads—includes some type of *inflection*. Angry, perplexed, sad, curious, excited, timid, etc., etc., etc.—no matter what the words are saying, their *inflection* almost always expresses some type of *emotion* when they are heard by us or spoken aloud. And, as discussed in the previous essay, in order to properly calibrate the inflection of dialogue according to emotion (something that likely involves Brocha's speech area) the dialogue obviously needs to be emotionally analyzed first. This means that our brain's executive decision-making mechanisms (which make *immediate* use of this emotional analysis) are likely set into inevitable motion slightly before that (*inflected)* dialogue can actually be heard by us.

Nonetheless, although our last ditch effort already looks dead in the water—in a final nod to free will's tenacious elusiveness—we'll explore one more full example of how consciously-experienced internal dialogue might *appear* to precede & activate an (inevitable) action. *(And we'll see if it really looks anything like the kind of free will you were hoping for.)* We can navigate this neural territory by considering a scenario: one of those *dodgy* married men is being heavily wooed by an attractive "other woman."

Observation of the woman's overtly & explicitly flirtatious behavior serves as raw data input that emerges from the right hemisphere as a pattern from which the left hemisphere constructs the internal dialogue: *"This chick wants me."* In the micro-moment before he hears himself say this—just after the Great Syntactic Divide—this man's emotional equations & beliefs might generate a combination of excitement & guilt that he *feels* as he has this thought. Has he decided if he's going to kiss her yet? Probably not quite. After this hypothetical observation is processed, the *next* thought might generate (among other neural items) the act of possibly kissing her. The enacting of this "kiss her" motor script might be inhibited by a number of factors beyond just the possible guilt, or the fear that his hypothesis is wrong. Matters of consent, appropriateness of setting, insecurity over one's kissing skills, a sudden craving for a cheeseburger, etc.—all can be factored in via sophisticated syntactically-based rule & vocabulary application, belief application, urge analysis, and emotional equations.

Let's say our potential couple is alone in a private room, and after she impatiently queries "*Are you going to kiss me or not?*" our married man finally *does* choose to kiss her. Did he choose *via free will* to violate his beliefs and kiss this clearly-non-wife-of-his? Well, let's say he *said to himself* "Okay—I'm just kissing her" in the moment *before* leaning in and planting one on her lips. And yet, if he was actually deciding to act *as* he was saying that to himself, he would (by the definition of action being ultimate proof of an *inevitable* decision) *already* be kissing her (*while* he was thinking this) and not just *talking to himself* about it (yet not actually *acting*). In fact, in the moment after having that thought and *before* actually kissing her he could *change his mind* and resist the impulse despite previously declaring internally his intention to act.

Thus—even if he consciously & linguistically expresses his "I'm just kissing her" *intent* in the moment *before* kissing her (and not *during*)—the moment in which he actually neurally *sets into motion* the "kiss her" motor script (his activation of that now-inevitable action) likely occurs just after the construction of the *subsequent* dialogue parcel. This means that his actual *choice* to kiss her still happens (as it *always does*) at that point just after all the necessary narrative-evaluating, decision-impacting data has been calculated & made available for use in choosing to activate or inhibit an action: within that DLPFC *executive* area that immediately follows all of those post-Great Syntactic Divide emotional & predictive calculations—and a couple steps before the accompanying dialogue arrives (fully-inflected) in our conscious awareness.

The conclusion seems inescapable: the neural *instruction* (or choice) to act is triggered via mechanisms that are likely enacted *just microscopically prior* to any consciously-experienced *awareness* of that choice is capable of being expressed or observed *via internal dialogue's* subsequent auditory-cortex-aided emergence in our Dynamic Core's multi-faceted arena.

Even if our married man skipped *all* of that internal dialogue foreplay & kissed her *immediately* in response to his first perception of her *wantingness* —the resultingly-triggered neural *instruction* to act would still depart from the same loop location & still occur *just microscopically prior* to hearing that perception's & action's "spurring" or *accompanying* thought. In reality, the *actual thought* that accompanies a moment of true action & decision like this is often less about *intention* than it is about *experience*: "This is awesome" or maybe simply "Oh my"—or probably something more along the lines of *"This is a bad idea."*

Damn You, Science

Even after exploring the hinterlands of our Great Syntactic Divide, it appears that science's verdict remains unchanged: we *decide* to act in the micro-moment *before* we *think* about the act (or have whatever thoughts that *appear to us* to happen in conjunction with, or somehow spark, choosing that action). This conclusion—drawn here from closely examining the looping neural mechanisms that we've proposed—is also supported by Benjamin Libet's famous research, which has detected within individuals apparent decision-making neural activity that precedes the subject's conscious awareness of choosing the action.[16]

Even though our brain truly does make choices *based upon* how we perceive ourselves to feel about the situation & our specific individual beliefs, our *awareness* of our brain's calculation of those factors in our decision-making process occurs slightly *after* our brain actually makes those calculations—calculations that are, in reality, what determines whether an action truly becomes *inevitable*. To which my ultimate response is, for the most part: *so what?* And I say this because we have yet to finally answer the *first* question we posed about free will, the philosophical one: *what truly is free will?*

In the end, when we humans express our desire to "have free will," we are essentially saying that we want to believe there is a true "Agent of Self" within us, an agent who *is* us, and through whom we consciously—and without any undetectable internal interference—control the choices, actions & behaviors that our body physically enacts. And I believe that our primary objection (or even *repulsion*) to the idea that decisions precede our conscious awareness of them is rooted in the belief that such a system robs us of that true Agent of Self. In other words, if the person that I *perceive myself to be* arrives onstage after the script has already been written, then my *perceived self* is merely an actor, and not a self-directed agent of any kind.

But what really *is* free will? Although there is no single-entity, fully-perceivable Agent of Self (it's actually our *whole system*—conscious & unconscious) we do, indeed, act as we feel we want to act—as long as we are *capable* of acting. Is this not essentially the core claim of free will? Does it matter whether or not we can consciously perceive *why* we are sometimes incapable of enacting our will? In some cases, a deeply desired or intended action is inhibited in ways that we are entirely conscious of (like a marathon runner who desperately intends to, but simply cannot, take that next step) and sometimes a deeply desired action might be inhibited in ways that we are not entirely conscious of (like desperately wanting to kiss someone who desperately wants you to kiss them, but simply not being able to act). In both cases we are still *aware* of our conflicting desires & capacity, and we feel that those factors ultimately reflect our inner *agency*—regardless of whether we actually can or do enact our self-expressed intention.

These mechanisms create a system that, *as a whole*, behaves exactly as any fully-autonomous Agent of Self (*with a sometimes limited capacity to enact its self-expressed will*) would behave. I like to refer to this conundrum—that we feel & behave exactly as if we had a genuinely autonomous, Agent-of-Self-driven free will, even though that solely-conscious, all-controlling Agent doesn't really exist—as the *Free Will Paradox*. This paradox is probably best expressed simply by the conclusion: *humans have, for all intents & purposes, genuine free will, except that they technically don't.*

Determinism-Schmerminism & The Truth About "Morality"

This brings us to one last philosophical notion that, like qualia, has held the spotlight for far too long in the arena of brain theory: *determinism*. Basically, determinism suggests that in the absence of something like a true "Agent of Self," all of the decisions & actions made by any creature with a purely physically-based system of mind would be *pre-determinable*—if you just happened to know *exactly* all the conditions that will contribute to this decision.

In other words, according to "determinism" *any* purely physically-based system of mind is ultimately a totally robotic & fully *agent-less* creature (since our choices are ultimately "merely" the result of at-least-momentarily "pre-determinable" neural & physical responses to our cornucopia of data input—as if that *miracle of existence* was something worth lamenting). Of course, this whole question of determinism has one fatal flaw: it has nothing to do with *reality*. In reality, there is absolutely no way that anyone could know *exactly* all the conditions that will contribute to a decision—which in many cases are virtually *uncountable* when you consider the myriad synaptic connections that are hit or just-missed in every round of thalamocortical processing (in every sensory, internal & cognitive system) and that contribute to every thought, action & *interaction* (which also exponentially increases result variables) that occur over a lifetime (which has been spent accumulating millions of unique data modules that also impact the predictability of those decisions).

The extraordinary neural complexity, malleability, "inter-causality" & "re-programmability" of this cognitive process are also central to the brilliant Peter Ulric Tse's anti-deterministic argument for free will in his book *The Neural Basis of Free Will: Criterial Causation* [17]. Ultimately, the entire *idea* of determinism is so uselessly irrelevant to anything that might relate to actually understanding or shaping human behavior that applying its principles to our *actual existence* is nothing less than totally absurd. The way that over-thinking (yet oddly short-sighted) philosophy-types try to make the question of determinism relevant to *actual human existence* is by saying crazy things like "if determinism governs all behavior, then we have no moral justification for punishing criminals, since their actions were not truly *chosen*, but merely the inevitable result of who they are & the situations they encountered." But this kind of logic is all twisted up in something else that philosophers talk a lot about, but that doesn't *really* exist: *morality*.

What the crazy statement above is trying to point out is that if determinism *does* truly rule the mind then punishing criminals is essentially *immoral* because they are not really *at fault* for their actions—therefore the criminal's actions themselves cannot be considered actually *immoral* (which is dubiously circular logic to begin with, since it means we also had *no choice* but to imprison them, making the whole question of whether we should or shouldn't moot). But *punishing* criminals for the innate *immortality* of their actions is *not* really why we imprison them. Essentially, we imprison them because—based on their actions—we predict that they are likely to act in this socially-destructive way in the future. Although throughout our cultures we talk about criminal systems being *punitive*—and founded upon an idea that criminal acts *deserve* to be punished—in the end, this is just our cultural & personally linguistic way of contemplating & expressing those deeply innate neural impulses to make decisions about protecting ourselves from future losses by assessing known evidence & applying reliable predictive patterns. (And as we mentioned in our previous essay, the notion that we use prisons for *reforming* offenders is simply a rather obvious *lie* that we pretend to believe because it makes us feel better about ourselves.)

In other words, despite what we say, we don't really imprison someone *just because they murdered someone & murder is immoral*. If we did, then there would be no such thing as *innocence due to self-defense* (or *innocence due to wearing some sort of official uniform while killing people*)—which we've conveniently declared as "moral" acts of murder. The real purpose, however, of such exceptions (to our brains) is to help sort out acts of murder that *are not* supposedly good predictors of future socially-destructive behavior, and thus *do not* require punishment to achieve a socially-desirable result.

Deep amongst our symbolic, contextual neural calculations—in the same way that a gain or loss is not about the *money* or *object*, but rather, how a change in our access to its *perceived* value will *help or harm* us—criminal punishment decisions are not based on judging the *morality* of the *act* of violence or theft, they're based on judging what those acts *predict* about *future behavior* that might help or harm us. And when certain societies or communities tend to punish, for example, specific races more harshly & frequently than the general population—even when committing *the exact same acts*—it's evidence that those decision-makers' minds are biased toward *predicting* that the "demonized" race is more likely to cause future societal harm. (Any individual's communally-nurtured brain-logic varies according to whatever data they've consumed.)

We've naturally built our system of societal rules & responses exactly how our brain works: by basing those rules & responses around making the best *predictions* about the *most desirable* or *most undesirable* future results (aka, the most valid & beneficial gain & loss predictions). From this perspective—as opposed to being a cause for *doubting* the reasoning behind imprisoning criminals—determinism is *exactly* why we should imprison people like violent criminals: because the predictability of behavior means that removing these individuals from general society is highly-likely to lead to an ultimately desirable result for that society, which is the primary goal of sequestering people like violence-prone individuals from the rest of us. (We also try to make everyone in society *aware* of potential punishments because we *predict* that the *fear* of such consequences will likely *prevent* at least some potential offenders from committing socially-harmful acts in the first place, aka, *deterrence*.)

Of course, because of those uncountable aspects of every decision, a *truly* determinism-based "Minority Report" & cognitively-predictive (somehow *neurally*-based or *gene-predisposition-*based) justice system is an entirely impossible *fantasy*. This means that we have to wait until someone *actually* does something awful to make a good prediction about whether or not they are likely to do something awful again in the future, and thus decide if they should be sequestered from the rest of us *apparently-much-less-likely-to-be-awful* humans.

Note that we didn't have to mention *morality* anywhere in *our* explanation of crime & punishment. Because there *is* no static or timeless *morality*. There are only culturally-developed, individually-learned *beliefs* that guide how we ultimately judge the "moral content" (aka, *social benefit*) of any act. And beliefs are not about any inherent *morality*—they're just a very special version of that thing our brain is obsessed with: a prediction. Beliefs are merely high-validity, high-value prediction tropes that help to guide our actions toward an ultimately desirable (gain-enhancing/loss-averting) result.

Morality, *per se*, is simply a non-starter where the brain is concerned—it just doesn't correlate to how our cognitive systems manage & judge data. All of those "moral" behaviors like aiding & sharing, affection, empathy, *not cheating* (on a test *or* your spouse), *forgiveness, etc., etc., etc.*—every apparently *self-sacrificing* or *purely-other-entity-benefitting* human action has been accounted for in our theory by some emotion, belief system or other *contextually-framed, survival-supporting, data-based & evolutionarily-arrived-at* neural mechanism. Of course, the ultimate societal *result* of applying all of these systems in a communal fashion over many millennia is exactly the same thing as what we consider to be a "moral code." Although—because the nature of our neural systems means that such "moral codes" are deeply & broadly culturally-based—the notion of morality is actually the *opposite* of what it's typically considered to be: in truth, *morality is highly malleable.*

Bring Me My Soul

In the end—even after dismissing determinism—Narrative Complexity's own kind of *paradoxical free will* is still not enough for most of us; we simply *want to be* that singular, fully self-informed & self-determining Agent. And as we stand on the deck of this once-wayward *Free Will* vessel now finally in the harbor, hollering into its empty hold for everything that it *has not*, at last, brought home to us—what we are really saying is very simple: we want a *soul*.

But what would a soul really be? Wouldn't the behavior & choices of a creature with a "soul" be the same as we behave & live right now? Can't this extraordinary, exquisitely-evolved & unimaginably complex system of mind be *equivalent* to a soul? Isn't each of our *minds* something that is uniquely us? Something purely based on a mix of unique inborn attributes, unique experiences, feelings, thoughts & desires, and uniquely acquired & organized rules, beliefs & vocabulary? Is this not, for all intents & purposes, exactly what a soul purports to be? Almost. But it does, of course, lack something very fundamental that is what we really seek from a soul: permanence. The human mind *cannot* give us permanence.

Forever does not appear to be something that these magnificent & genuinely *soulful* neural mechanisms & systems can give you. And, frankly, that pisses me off. I *want* some kind of permanence. I want *to see what happens* and *be part of it*. I *am not* "okay" with my non-permanence just because I won't actually be able to *perceive for myself* the horror of my non-consciousness. I am utterly terrified by a state of being that I will never know.

And so, for me, this is what I have: my *desire* to be here—which seems like an ultimately irrational (or at least overly-circularly-logical) motivation for being. *Be to be.* And yet I *believe in* and *cling to* that desire to be here. For me, this fervent, life-defining & ultimately-irrational desire is the closest thing *that I have* to a soul. It *is* me. The thing that says *I am* and *I want to be*. The thing that someday will be *was*—a thing that I wish would be capable, in that someday, of saying: *"I was. I was."*

Direly, based upon everything I have so far learned in this life, and all of those self-defining beliefs, and rules, and words, and experiences—I simply do not believe it is likely that this mind will ever someday say to itself, *"I was."* We live, I believe, by definition, in the universe of the present—it is the only place we ever truly are, or will be. We exist now.

And I believe that in all the ways that might genuinely matter—we, the unique being & mind that is each of us, *do* have free will & full domain over the choices that we *perceive* to be perpetually presented to us. It is simply that a great deal of that unique being & mind is making its contributions to our ongoing self *behind the curtain*. But the work that goes on unseen is just as much a part of *who we are*—that unique amalgam of self-accumulated & self-organized data—as who we consciously *perceive* ourselves to be.

I am here. We exist now. You are a mind in the present. The most extraordinary expression of *self* that this earth has ever created. You will know nothing but this, but you may try to know as much of this as you might desire while you are here. Within the confines of your circumstances, you may choose to do with being here whatever you wish—even to quixotically battle those confines, to seek to alter the world in which you roam. That freedom, this mind, its temporariness, and the will to do, and be—these are what we have been given.

The will of the free, and a mind for the now. A place we are in time. The melancholy glory of being.

###

Endnotes:

p. 204:
1. Bargh, John A., and Ezequiel Morsella. "The unconscious mind." *Perspectives on psychological science* 3.1 (2008): 73-79.

2. Morsella, Ezequiel, et al. "Passive frame theory: A new synthesis." (2016).

3. Edelman, Gerald M., and Giulio Tononi. A universe of consciousness: How matter becomes imagination. Basic books, 2000.

p. 205:
4. Deacon, Terrence. The Symbolic Species. Norton, 1999.

p. 209:
5. Wymbs, Nicholas F., et al. "Differential recruitment of the sensorimotor putamen and frontoparietal cortex during motor chunking in humans." *Neuron* 74.5 (2012): 936-946.

p. 214:
6. Domenech, Philippe, and Etienne Koechlin. "Executive control and decision-making in the prefrontal cortex." *Current Opinion in Behavioral Sciences* 1 (2015): 101-106.

p. 216:
7. Rolls, Edmund T. *Emotion and decision making explained*. Oxford University Press, 2014.

p. 223:
8. Klein, Johannes C., et al. "Topography of connections between human prefrontal cortex and mediodorsal thalamus studied with diffusion tractography." *Neuroimage* 51.2 (2010): 555-564.

9. Alexander, Garrett E., Mahlon R. DeLong, and Peter L. Strick. "Parallel organization of functionally segregated circuits linking basal ganglia and cortex." *Annual review of neuroscience* 9.1 (1986): 357-381.

10. Kultas–Ilinsky, Kristy, Elena Sivan–Loukianova, and Igor A. Ilinsky. "Reevaluation of the primary motor cortex connections with the thalamus in primates." *Journal of Comparative Neurology* 457.2 (2003): 133-158.

11. Marsden, C. D. "What do the basal ganglia tell premotor cortical areas." *Motor areas of the cerebral cortex* (1987): 282-300.

p. 225:
12. Andersen, Richard A., and Christopher A. Buneo. "Intentional maps in posterior parietal cortex." *Annual review of neuroscience* 25.1 (2002): 189-220.

p. 226:
13. Haller, Matar, et al. "Persistent neuronal activity in human prefrontal cortex links perception and action." *Nature human behaviour* 2.1 (2018): 80.

p. 228:
14. Kahneman, Daniel, and Amos Tversky. "Prospect theory: An analysis of decision under risk." *Econometrica: Journal of the Econometric Society* (1979): 263-291.

15. Tversky, Amos, and Daniel Kahneman. "Advances in prospect theory: Cumulative representation of uncertainty." *Journal of Risk and uncertainty* 5.4 (1992): 297-323.

p.230:
16. Libet, Benjamin, et al. "Time of conscious intention to act in relation to onset of cerebral activity (readiness-potential) the unconscious initiation of a freely voluntary act." *Brain* 106.3 (1983): 623-642.

p.232
17. Tse, Peter Ulric. The Neural Basis of Free Will: Criterial Causation. MIT Press, 2013.

A Comedic Addendum | *Humor*

The Need For Novelty *or* Why Is Stuff Funny?

Humans do not *all* find the same things to be funny, but all humans do find *some* things to be funny. The ability to detect humor & the tendency to reflect (through a smirk or a laugh) the detection of that humor are universal human traits. And if our brain has turned these mechanisms into universal human traits, then humor *must* have some purpose beyond simply adding a little entertainment to our lives. This conclusion leads us to a couple of obvious questions: why is certain stuff funny, and why do our brains care?

What Is Humor?

First, we have to define exactly what human brains judge as "funny." In terms of our *response* to amusing stimuli, that's a simple task. The brain mechanisms that are engaged by our humor response have a fairly obvious *tell*: they cause us to *physically express* our amusement somehow (often unexpectedly or even uncontrollably). This expression of amusement occurs along a broad continuum —ranging from that mere smirk to hysterical fits of laughter. No matter where along the spectrum one's response falls, if our brain has identified something as funny (or even just *quirky*), we're highly prone to show it somehow in our face.

What, then, is our brain actually *identifying* when it deems something as funny or quirky? *Novelty*. Every joke is, in essence, a *surprise*. Whenever you smirk or laugh or are otherwise *amused*, your brain is tagging that surprising event, observation or narrative as *uniquely novel*. Why does our brain care so much about novelty that it's devised a special universal human mechanism devoted to identifying & analyzing uniquely novel data? Because above all else, the human mind & consciousness are built to maximize that primary, evolution-conquering tool: creative problem-solving.

In the brain's game of creative problem-solving, novel data & patterns *always* have some potential future value. Whether it's the smile-producing & *genuinely odd* way that the errantly-floating feather seemed to skip alongside your feet (and whose uniquely new presentation of locomotion might spur an engineering *a-ha*) to that *hilarious* narrative twist you didn't see coming (but which, since it's now been experienced, you can apply as a possible narrative predictor in a future unique circumstance).

This creative problem-solving boost is the same value we get from clever word-play jokes like puns & other novel verbal linguistics—which are the source a literally *volumes* of humor. As we first explained in essay #1, the human brain's thought-conjuring & problem-solving machinery is primarily language-based, and the ability to cross-associate unlike ideas via their related modular word-based components is *enhanced* by words that have those unique, highly-malleable & flexible multiple-associations. Puns & word-play jokes help reveal to our brains even the most-hidden of these useful multiple-usages & associations between words, which supports humor's ultimate purpose: to aid in creative problem-solving.

From our mind's point of view, every current or upcoming unsolved problem might be unlocked someday by that as-yet-un-encountered, uniquely-novel data or pattern. For these reasons, the human brain is built to *love* consuming & cross-applying novel data in its quest for more creative & unique solutions. And what does our brain do when it wants us to love something? *It feeds us pleasure in response to that something.* Our brains need novelty, therefore, humans love humor.

Parsing The Pleasure

This pleasure response mainly does two things for humans: it helps us to take special note of the uniquely novel data or pattern, and it helps us to *remember* the novelty for future use. This is why it's so useful (& adorable) for babies to be giggling & smiling at stimuli all of the time. And to a brand-new baby, nearly everything that's even the slightest bit novel is likely to be judged as uniquely (and thus, humorously) novel.

Particularly in those earliest (infant & toddler) stages of our brains' cognitive-rule development processes (mechanisms thoroughly explained in essay #4)—discerning, distinguishing & remembering specific new patterns is vital to building & sorting the plethora of new rules that our left hemisphere is stocking for a lifetime of use. As we age, the needs of our brain change, leading to a change in the way we respond to that uniquely novel data (aka, our *sense of humor*).

For one thing, you might say that our "giggliness-quotient" decreases with age. That once unendingly-amused toddler eventually, over the years, tends to find fewer & fewer events, observations & narratives uniquely novel. Although the *full scope* of what we find to be funny generally grows much broader & deeper as we age, it seems that the sheer number of experiences that we judge as humorous is still much greater in our youth.

However, in terms of humor & aging, what we give up in quantity might be made up for in quality. Mature, experienced & fully-analytical minds are likely to find humorous experiences more *rewarding*. This is because, generally speaking, those minds are ferreting more-useful data out of these novelty-based experiences. How exactly does our brain turn a novel experience (aka *a comedic narrative*) into more-useful data? By using the same primary tools that it uses to analyze all narratives: the value & validity judgements that fuel our emotional equations (explored extensively in essay #2).

According to Narrative Complexity's hypothesis, the human brain makes 7 specific narratively-based judgements when analyzing any of these uniquely novel events or comedic narratives. In other words, there are 7 judgements that the brain makes when determining its response to a joke. As we age & our minds mature, our brains tend to weight some of those specific judgements differently, leading to both different responses to novelty & different preferences for certain kinds of uniquely novel experiences or narratives.

When we're younger, our brains tend to be less capable of flexibly & subtly balancing & analyzing those 7 judgements of a comedic narrative, and are thus less capable of receiving the humor-based pleasure derived from narratives that generate their comedy in more balanced & complex ways. Younger individuals have what we might think of as a less-sophisticated humor palette, one that only requires high scores in a few key categories in order to generate our humor response, but that is less responsive when those categories are more balanced with the other 7.

These younger humor palettes are also willing to soak up lots of seemingly-redundant experiences that hit those few key notes—because they're likely still developing their initial sensitivities to the subtleties *within* those few key categories. In the construction of the human mind, one of the brain's complexity-developing tricks is to start narrow & go deep—providing complex-but-microcosmic early neural models to found the building of broader, more robust & more flexible mechanisms later. In terms of humor, this means that in order to teach your brain how to eventually laugh at a *New Yorker cartoon*, in your youth you must first master an understanding of all the subtleties of *fart jokes*.

What exactly are the 7 independent, narratively-based judgements that our brain makes about uniquely-novel data? Behold, the anatomy of a joke:

1. Likelihood Judgement - This first judgement is essentially about the expectation or the "set-up" of a joke; therefore, this defines the general emotional state of our mind *prior* to encountering the joke's surprise.

During a comedic narrative or event, we either see a novel twist coming or we don't. If we feel that there is a strong likelihood of something unexpected happening (a joke with a "set-up") we feel the *anxiety* of that predictive uncertainty. If we are *not* anticipating something unexpected happening (low likelihood or no "set-up") then we feel the security of our predictive *confidence* (and are thus very likely to be more surprised or even *shocked* by the upcoming novel twist).

2. Loss/Gain Judgement - Once the unexpected event occurs (in essence, the "punchline") the first thing our brain does is determine whether the event represents a loss or a gain to us. If the surprise is a rock falling on someone's head, that's likely viewed empathically as a loss (making it a pain-based joke). If the surprise is a diamond necklace falling on someone's head, that's likely viewed empathically as a gain (making it a pleasure-based joke).

However, this judgement is also impacted by how we feel about the individual to whom these events happen. If it happens to someone whom we don't like or whom we have disdain for, then the rock might feel like a gain & the diamond necklace like a loss. No matter how many different factors are at play here, the unexpected event is ultimately judged by our brain as a personally-felt gain or loss.

3, 4 & 5. Importance, Relevance & Novelty Judgements - These three judgements (which are essentially simultaneous) are those fundamental measurements that the brain uses to determine the overall *value* of the loss/gain identified in judgement #2 above.

If the rock that unexpectedly falls on the character accidentally kills him (or if the falling necklace came from the Titanic) that makes the event more *important* (and the joke more *outrageous*) than if the falling rock merely annoys him. If the character is a small child and you also have a small child, that might make the uniquely novel event more *relevant* (essentially making the joke more *insightful* to you). And if you've never unexpectedly seen a rock fall on anyone's head before (because you've been living under one—or maybe you're, like, 2) then this event might actually seem highly novel (increasing the joke's most vital element, its novelty-based *humor*).

Together these 3 judgements essentially determine the *intensity* of our emotional/

physical response to the joke. Low scores across these categories create a smirk; high scores in these 3 are likely good for big laughs.

6. Reliability Judgement - After observing (or experiencing) the unexpected & novel result within a comedic narrative (and feeling those initial emotions) our brain wants to assess the actual *usefulness* (or the *impact*) of this attention-grabbing new data discovery.

In other words, some of these surprises provide data that's more valid—more reliable as a behavioral or narrative predictor or model in the future. Although *all* punchlines are unexpected or unlikely narrative results, the *plausibility* or the ultimate *truth* (to us) of that unexpected result (often gleaned after a moment of post-surprise reflection) helps to determine our different subsequent validity-based emotional responses to the punchline.

If, in the end, the surprise feels contrived or phony—making it more unreliable as a predictor—this tends to dampen our enthusiasm for the joke. In contrast, if the punchline or unexpected twist feels *especially* true or plausible—declaring itself a reliable predictor—that tends to bolster our enthusiasm for (and the pleasure derived from) the joke.

7. Belief Judgement - The other half of this post-surprise *assessment* of a novel result's usefulness or impact: determining whether the comedic narrative or event complies with or violates any of our beliefs (defined in essay #2 & explored further in essay #4).

In the end, even if we initially (and somewhat involuntarily) laughed at a joke, during this assessment our smile might still morph into an expression of *disgust* if the punchline or character behavior ultimately *violates* one of our stronger beliefs. Comedic narratives are, after all, still *narratives*, which means that (according to Narrative Complexity's mechanisms) before they enter our conscious awareness they're automatically analyzed by our belief system for emotional generation.

And when a comedic narrative scores high in belief *compliance*, it tends to enhance our *connection* to the humor & its source—a result of that admiration-based modeling mechanic triggered by others who demonstrate compliance to our beliefs. Comics that play heavily with these belief judgements are the kinds of comics who tend to inspire devoted worship: individuals whose comedy is founded upon strong & distinct beliefs that are shared by its audience. These are the controversial, boundary-pushing & revered comedians (like Joan Rivers, George Carlin, Richard Pryor, Bill Hicks, Marc Maron, Chris Rock, Doug Stanhope, Louis C.K., Sarah Silverman, Tig Notaro, Dave Chapelle, and Hannah Gadsby) who not only seem to speak uniquely novel & cleverly arrived-at high-value truths, but surprisingly *profound*, belief-defining truths.

In fact, challenging-but-worshipped comics like those named above (or the ultimate example: Lenny Bruce) tend to build their envelope-pushing comedy around a clever belief-engaging trick. These comedians usually work off the premise that "the truth rules above all"—this belief posits that nothing is more valuable than expressing the truth, even if it is offensive or painful. Then they reveal the most offensive or painful truths that they can muster, but do so while complicatedly & cleverly remaining within the confines of "truth-telling" about some high-value topic, allowing (or forcing) us to "accept" the offensive or painful (yet still important, relevant *& hilarious*) unexpected truth.

In a comedic situation or narrative, tolerating the *violation* of a powerful belief in the service of complying with an *even higher* belief causes some *very interesting* (& often oddly pleasurable) emotional responses in humans. Part of what we're feeling is likely the result of little neural renovation, because jokes like this probably cause some subtle rearrangement of our own belief structures (in order to accommodate this clever new comedic conundrum).

In other words—no matter how it's structured—deep down in our brains, a joke is rarely *just a joke*.

The Comedy Gun
Before we go, I'll leave you with a little eye candy. In order to provide a more visual way to break down his 7-step comedic process, I've built a handy chart—something that's a bit like the mutant offspring of the *Mothership of Emotions* (presented by our theory in essay #2). And this mutant offspring has its own semi-clever name: *The Comedy Gun* (a tiny homage to that deathly classic, and truly-certainly-never-funny-to-begin-with comedy "device" of prop guns that either explode loudly or spit out a silent, dangling "Bang!" flag)...

the Comedy Gun

the Lead	the Surprise				the Assessment	
Likelihood Judgement	Loss/Gain Judgement	Value Judgements			Reliability Judgement	Belief Judgement
Strong Anticipation of Surprise >Set-up< (Anxiety)	Loss (Disappointment)	Low Importance	Low Relevance	Low Novelty	Unreliable (Invalid Pattern)	Belief Violation (Disgust / Guilt)
(For each element of comedic narrative, judgement either ^ ~ v . Each element judged independently; 128 combinations.)						
Weak Anticipation of Surprise >No Set-up< (Confidence)	Gain (Delight)	High Importance	High Relevance	High Novelty	Reliable (Valid Pattern)	Belief Compliance (Pride / Satisfaction)
Expectation	Pain/Pleasure	Outrageousness	Insightfulness	Humor	Truth	Connection
Barrel	Intensity				Impact	

A Hard Problem Addendum |

Experiencing Existence

Why The Hard Problem Doesn't Really Exist *or* Why Do We Humans Experience Our Existence?

Something Fishy

I've got news that's not really news: you *experience* your consciousness. Those myriad sights, sounds, tastes, scents, objects, entities, environments, thoughts, and *feelings* that our consciousness integrates (via our unified & fluid "Dynamic Core" of awareness) are all perceivable to (& reportable by) *you*, or more specifically, to (& by) your brain—that vast collection of neurons that actually *presents* all of those elements to other neurons within that same collection. In addition, as you have likely noticed, we experience having a body that is *within* that larger world, yet is still distinctly *separate from* everything else in that world. Together, all of these truths are more commonly described as being a living and awake human.

Alas, although it is the most self-evidently obvious aspect of consciousness (that we experience our existence) some people think *there's just something a little fishy about it*. (Actually, due to our ancestral neural debt to lampreys, there literally *is* something *fishy* about our conscious experience.) And if you spend a lot of time thinking about consciousness (which I'm deeply guilty of) the matter can become more elusive:

I mean, yeah, obviously I'm experiencing all this, but...why, and how? Why shouldn't this machine in my head just pass along all of this data to all those cognitive processes without "presenting" it to me? And who is "me" anyway? If the brain is merely passing along this data to itself & its own cognitive systems, where do "I" come in to the picture? Who is the brain presenting this to if there is no "watcher" (or "feeler") in the system? And if this "watcher/feeler" is not necessary to the system—then what am "I" and why do "I" exist to experience any of this at all?

A-ha! Maybe there's something else happening here in the brain, something beyond just neurons & synapses—some property or mechanism within these systems by which I (this watcher/feeler) emerge to experience this consciousness as this data is being integrated & passed along to our cognitive systems. What is really going on here? How can a purely physical system like the brain perceive the experience of its own existence, and why is it bothering to do this at all—isn't "being" more than just some incidental & ultimately-illusory emergent property of all these systems processing all this data?

That last multi-pronged question that someone gets to at the end of that fishy slippery slope, that's essentially what philosopher David Chalmers in the 1990s deemed "The Hard Problem" of consciousness—the problem that supposedly simply eludes any plausible possible solution within the confines of current neuroscience. However, from Narrative Complexity's point of view (and one of the reasons why we didn't directly address this matter in the main essays) once you piece together all of the functions & purposes of consciousness' main mechanisms, the real question is: why *wouldn't* a machine like the human brain experience a perception of its own existence?

In other words, as we'll show here—per the way that vertebrate brains primarily function—one of the machine's *main purposes* is to generate a specifically-depicted, fluid & unified model of the creature & the world around it, a model that *must necessarily* be perceived by & handled by (aka, *experienced by*) that same machine's own cognitive & behavior-generating mechanisms *as a specifically-depicted, fluid & unified model*. In fact, chordate evolution clearly didn't find "The Hard Problem" to be very hard at all, and it went about solving the matter fairly early on—right about the same time that it came up with *spines*...

Vertebrates vs. Robots

We should start here by offering an apology to all of those pre-mammal vertebrate minds that we referred to in Essay #1 as *robots*. That was probably a little harsh. Although, yes, their entirely pre-programmed & reflexive action responses make those creatures pretty *robotic* in their behavior, they *do* feature a capacity that is pretty *non-robotic*: they likely *experience* their existence in much the same way that humans & other mammals do. This is because (beginning with those lampreys and their unique neural arena integrating incoming visual & electrical data) they also likely possess a more rudimentary, but similarly unified & fluid "Dynamic Core" of awareness that produces a specific internal model of their world & bodies by integrating widely-varying sources of environmental & physical sensory input.

The *real* robots among us—creatures whose entirely pre-programmed & reflexive responses are *not* likely directed via one of those unified & fluid models produced by integrating various sensory input—are actually *insects*. (This category of robots also includes *in*vertebrate chordates like sea squirts & hagfish—but insects can display much more interesting & *conscious-like* behavior than our invertebrate ancestors, so we'll focus of them here.) If you wanted to propose that experiencing ourselves & our world is a central & necessary element of any kind of true consciousness, then you would have to declare that insects are not conscious entities. (A 2016 paper from Colin Klein & Andrew Barron disagrees, but we'll address their hypothesis in this essay's post-script.)

Why haven't insects developed their own consciousness? Probably because they've always been able to effectively go about their daily business without creating an *internal model* of the world around them. Sure, the most sophisticated insects—like bees—create awesome data resources like an internal *map* of their paths between the hive and food sources. It's been found that bees even use visual landmarks to help guide themselves during this process. Nonetheless, for reasons we'll explain here, they are *not* likely integrating any of that data into a unified, fluid *model* of their environment.

In bees, that spike visual data (e.g., something large or brightly-colored) likely simply triggers some kind of *path-location-via-association* mechanic within that internal map, which can help to direct & maintain motor scripts that were initially triggered when they left the hive or the food source. In other words, they aren't "seeing" this map in their little heads and overlaying the path on a fluid visual depiction of their environment, then using that overlay to guide their flight within that modeled environment (like some kind of insect Terminator). When they first found the food source, they merely saved their internal "this far, that direction, etc." path-map and then used that same data to direct their return to the hive. On the way there or back, they might also associate some spike visual data with a specific location in that path-map that helps to re-orient them during future food runs if they drift off-course.

A system like this simply doesn't require that visual data and map data to be perceived as a unified model, because it is neither tracking distinct objects within that environment (which requires a neural process dubbed "binding") nor guiding actions according to how those objects & environment are presented *within* any kind of internal depiction. Thus, a bee's visual data & path data can be fed directly into associative & behavior-generating processing systems without any need to create a complex model that must be perceived & responded to *as a model* in order to guide behavior. A bee can do everything it needs to do in essentially this same fashion—*no conscious experience necessary.*

This is why, even though a bee or a fly or moth has eyes and may seem to be seeking its way out of a room with an open window, it might endlessly zip around that room randomly, unable to escape despite the clear presence of an obvious escape route. Ultimately, the insect isn't really *aware* of anything around it because it is not using that sensory data to construct a unified model. In a case like this, the insect is just running a rote flying motor script until raw sensory conditions trigger a different response—like if you darkened the room then turned on a lamp, directing its flight toward the source of that new raw & highly-defined (spike) sensory data.

Even when *bee-*havior (sorry, irresistible) appears to involve tracking distinct objects in a way that might require sophisticated on-the-fly visual modeling, simpler pre-programmed (& *non-dynamic*) mechanisms can still easily do the job. Consider the *waggle dance*—a specific (but highly-variable) figure-8 flying pattern that bees use to direct other bees to food sources. In these cases, the "dancing bee" visual data is likely pre-programmed to be recognized & treated as *special case* visual data that automatically engages with specific systems like navigation. Basically, when returning from a new food source, bees are programmed to display specific flying behavior (generated via their just-recorded map data) that, in turn, directly programs other bees navigational systems via the processing of that *special case* visual data. In other words, this unique bee-to-bee visual programming language comes built into their neural systems (in the same way that their bee-to-bee olfactory programming language is built-in). Thus, this is not a case of gathering & processing new visual data and *dynamically* modeling distinct, representative (& trackable) objects—the highly-specific *special case* visual data input merely serves as a source of easily-identified programming code.

Similarly, even in situations that *seem* to involve bees identifying goal-related targets, their robotic natures can still be seen beneath their actions. For example, if a bee hive is violently disturbed, this sudden displacement within the hive might trigger threat-response "swarming" or "attacking" behavior in the bees. Although a quick response means that the bees will likely swarm the actual disturbing party—*who* they swarm is irrelevant to the bees (& not even on their tiny minds). They simply fly together and respond to each other's cues (via *special-case-data* code) until some of them encounter some appropriately "spike-ish" data: a warm, soft surface or supple, sting-able flesh. And once they've associated this new data with their attack scripts (like the specific color of this sting-able surface), they can continue to direct those scripts via that data—aka, pursue their newfound target (at least until the associated-data-possessing entity can get enough distance or cover to effectively disappear from their pursuant's raw-visual-data radar).

Furthermore, when encountering their tormentor, the bees will not make any judgement about whether the entity is highly-formidable or easily-conquered, and thus (unlike a large swath of their vertebrate counterparts) they cannot determine whether to swarm or retreat based on judging the most-likely-to-be-beneficial result (unless that tormentor, say, unleashes a flamethrower and that raw & highly-defined sensory data directs them to fly oppositely). This is because basically, to the insect, *there is no entity*—there's just raw data & reflexive responses.

In truth, to the insect, *there is no insect*. Because they can do everything they need to do simply by using those robot-like systems —*no conscious experience necessary*.

It's All About The Model & All In Your Head

You might've noticed that I've been making excessive use of the word *model*. And it might be getting on your nerves a little… *I get it, our brains make a model of the world & insects' brains don't, but what's really your point here—aren't insects still experiencing the world via their sensory organs in some kind of fragmented state?* That seems difficult to answer without being able to actually *ask* an insect what it's experiencing (unfortunately, "reportable-ness" is still a key element of consciousness research). However, if we apply some Occam's razor logic and presume that insect wiring would most likely only achieve the minimum amount of sophistication required to carry out reliable, adaptive behavior—then it would make most sense for insect behavioral systems to handle raw data directly, without any intermediate "present & perceive" processing that specifically reconfigures the data before it's received by those systems.

This is why I keep trying to drive the word *model* into your head (even though, technically, that model is *already* deep inside it). Because if you don't require a model to carry out your behavior (like the system we just described in bees) then there is no need for any kind of mechanism that might qualify as "perceiving" (aka, *experiencing*) that sensory data. In essence, those non-conscious insect behavioral systems are merely *interacting* with that raw sensory data and responding accordingly.

What reason might we have to suspect that this kind of interaction between raw sensory data & responding neural systems is not experienced simply because it is not being processed via some model creation/perception mechanism? *Because we process data all the time in this fashion without experiencing it*. Our brains are receiving &

responding to all kinds of incoming data without our conscious perception of the process. And in all of these cases, the reason that we don't perceive the process is ultimately the same: because the data did not get subsumed by our internal model's attention-driven perceptual processing. This incoming data was either never integrated into that internal model, or (despite being integrated into the model) it was essentially ignored or discarded by our perceptual processing because it did not garner enough of our (limited) attention.

For example, our brain receives all kinds of internal data about our bodies and responds appropriately (to regulate systems like digestion & respiration & on & on) without our conscious perception of those processes. That's because most of this data is never fully integrated into that model of our world & ourselves. At best we might perceive some internal pains in some places or some *gurgliness* in our guts—but the actual physical internal undulations of that gut as it produces the *gurgliness* are totally invisible within our model, and thus, are outside our capacity to truly perceive or experience them. (Consider: if your arm suddenly started undulating like your intestines, you'd *definitely* perceive it—ultimately, you've got no idea what's really going on inside you, and you probably don't want one.)

In addition, there are those countless cases in which data that's integrated into your internal model simply garners none of your attention—resulting in you "not seeing" something that is otherwise clearly & obviously there. In reality, your brain received & (at least *partly*) integrated that visual data into your model, but the lack of attention kept it from being subsumed by those perceptual processes that allow it to be "seen" via our consciousness. This kind of perceptual mechanic was famously demonstrated in an experiment in which viewers of a basketball game (who were instructed to count the number of passes, thus focusing their attention) failed to perceive a person in a gorilla suit running by in the background. Oppositely, our total dependence on this model for perceiving the world is also illustrated by those myriad objects & entities that you *think* you see, but are really just *imagined additions* to the model (thanks to things like overeager predictive/sensory mechanisms & neural dysfunction).

A well-trained skeptic might see a small opening here for some kind of limited insect experiential consciousness: *So… what you're really saying is that the subsuming of "attention-defined" data—which is essentially a version of "spike" data—allows a creature to "experience" that data. Why, then, wouldn't insect brains experience the "spike" raw (non-modeled) sensory data that they subsume?* For starters, as noted, in humans our own high-priority (spike) internal data still goes unexperienced *if it isn't integrated into the model first.* Nonetheless, even if we accepted that insects "experienced" that subsumed spike raw data, their conscious experience would be *highly* intermittent & fragmented: stretches of nothing that are occasionally interrupted by quick flashes of brightly-colored images or sequences of special visual or olfactory data code input. Not only would such a conscious experience be completely unlike vertebrates' fluid, unified conscious experience (produced by the *constant & ongoing* subsuming of integrated, widely-varied sources of attention-defined/spike data) but such a conscious experience would be a totally useless & extraneous addition to those insect cognitive systems. As we'll explain shortly, vertebrate conscious experience (& mammalian consciousness in particular) is both a *deeply useful* & *innately necessary* element of those creatures' cognitive & behavioral processes.

Collectively, this all tells us a good deal about the core functions & purposes of this model—and its necessary role in any conscious experience. For one, it tells us that the model isn't primarily meant to help out with or govern any of those internal systems. Most of that internal data is routed directly (& imperceivably) through other systems. It also tells us that in order for data within this model to actually initiate conscious actions & responses, merely being present isn't enough—that data must also be processed by those attention-defined perceptual mechanisms. As demonstrated in Essay #5's description of reaching for the wrong cabinet, even actions that are initiated via our conscious arena can slip into that "action-maintenance" mode and be directed entirely via unconscious (& unperceived) data intake & response processes.

If something isn't presented & *perceived* via our internal model, we aren't really conscious of that something. This is why such a model plays a necessary role in any conscious experience—because this model basically *is* our consciousness. In truth, our cognitive systems aren't actually interacting directly with the "real world" in the same way that those robotic insects are via their raw-sensory-data-straight-to-response-processing systems. Vertebrate cognition is primarily interacting with the model of the world that is being created & sustained *within our minds.* Thus, even when something exists in the "real world" and is technically within our purview, if our mind doesn't integrate it into our model we're totally unaware of its presence. (There are, for example, all kinds of radiations & light waves, etc. that are absolutely in our presence, but completely unperceived via our model.)

The world that we live in—the universe that we *experience*—is all about the model, and it's all in our heads. *That's* the reason why I keep talking about this model (and that's *what's really my point here*).

A Machine Is As A Machine Does

Hopefully our discussion thus far has, at least, reduced the matter at hand to "The Rapidly-Diminishing Problem." But there are surely still some Hard Core Hard Problemites who have some very deep concerns here… *So what? I mean, really, so what? So some model gets created & perceived. All you've done is sneakily try to treat the words "perceived" & "experienced" like synonyms and declared the matter mostly solved. That's nonsense! Why must perception equal experience? The Hard Problem isn't simply some mix & match word game, it's about telling me what happens in the brain that turns that perception <u>into</u> experience, and explaining why this perception can't just go about its neural business without generating that experience. You're wasting my time.*

That might be a little harsh, but the Hard Core Hard Problemites have a point—*something still seems a little fishy here.* However, with one last (& kind of reverse) slice of Occam's razor, I think we can get to the heart of these final objections. Earlier I'd said that after examining the functions & purposes of consciousness' main mechanisms, the real question is: why *wouldn't* a machine like the human brain experience a perception of its own existence? Consider that The Hard Problem is based on the premise that the functionings of our highest cognitive systems like language & beliefs (the domain of our main essays) are all the *easy problems*—suggesting that conscious experience *must* involve more highly-complex mechanisms. But evolution shows us that the most complicated neural mechanisms are almost always *preceded* by those "easier" systems. And evolution also shows us that all of those "easier" systems like language & beliefs were actually preceded by the development of those experiential modeling systems *by a few hundred million years*. Why should a problem that was likely handled by evolution in *lampreys* require some solution more magical or unimaginable than the solutions that resulted in *human language*?

This is what I mean when I say that The Hard Problem doesn't really exist: because nothing in our understanding of the universe suggests it's *more likely* that a machine designed to create a highly-specific experiential model of its world & body—in order to guide actions according to how specific data within that model is subsumed—*would not* subsequently & necessarily experience the perception of this model as that data is processed. *That's the whole point of the model—for the machine to perceive & experience the world via subsuming data within this model.* "Experience" is the process of the machine perceiving data that the machine itself has modeled specifically to enable the process of perception—*and we are that machine*. To top it all off, human brains have added their own singularly-unique element to this model, one that allows mere *experience* to transform into *self-awareness*: internal dialogue.

The reason why *feelings* (aka, emotions & physical sensations) are included in that experience is because those aspects of our consciousness include valuable data about our world & ourselves—data that is necessary for creating associations & directing conscious behavior, and thus, data that must be integrated into that model in order to help accomplish those tasks. And it is partly the looping nature of our consciousness that makes the integration of feelings into this model necessary to functioning—even in pre-language mammals that are incapable of consciously *thinking* about those feelings.

Consider that when a puppy is groomed by a caretaker (canine, human, *ad infinitum*) those physical sensations (experienced within the model) are subsumed by cognitive processes that respond (in part) by generating pleasure-producing oxytocin. One of the purposes of that oxytocin is to aid in *bonding* that puppy with its caretaker—enabling more trusting behavior between them in the future. The "pleasure-sensation" that the puppy subsequently experiences can then become *associated with* whatever entity is simultaneously present within the model & garnering the puppy's attention. In addition—because those pleasure sensations are focused on (or most concentrated within) the bodily area where the grooming is actually occurring—those feelings also help draw specific attention to whatever entity is enacting the grooming.

Importantly, it is the subsuming of these *second-round* interactions within the model that allows the entity to be properly associated with that pleasure & encoded within the cognitive systems as "trustworthy." Although the first interaction that was subsumed via the model (the physical sensations of grooming) is ultimately the root source of the association & encoding, that data alone cannot do the job—it needs to generate the emotional response first, and then that pleasure *must necessarily* be experienced within that model in order for the actual bonding to subsequently be encoded within the cognitive systems.

Again, our well-trained skeptic might interrupt here—this time questioning the real *necessity* of those second-round interactions: *Why isn't the first interaction enough data to do the job? Why can't the grooming automatically trigger a positive association to the groomer as soon as those physical sensations are generated?* Because those physical sensations can't trigger any cognitive or emotional response (like bonding) until they're actually subsumed by our perception. And once this data has been subsumed & triggers a bonding response, it's winding through that unaware, *unconscious* cognitive maze—which means that cognitive/emotional response has no immediate access to any of the modeled sensory data required to complete the bonding task. *This* means that bonding response must subsequently emerge somehow *within that model* in order for it to be associated with (and help direct attention toward) whatever entity is simultaneously present within the model & enacting the grooming.

How do vertebrate brains represent that bonding response within the model? *Feelings* (which are mapped to that modeled body). How else is the brain going to represent this data? Should it turn the visual arena a special color—a color that triggers bonding within our cognitive systems when simultaneously subsumed with specific other-entity data? Should the response be represented by auditory data, or olfactory? Any of those might actually do the job, but as we discussed in our emotions essay—*all* of our feelings are ultimately rooted in those ancient pain & pleasure responses, responses that were *all* related to *bodily* matters like hunger & injury. Thus, those pain & pleasure responses were represented in the model via bodily-based sensations (aka, *feelings*)—and all subsequent pain-&-pleasure-rooted, feelings-based data has continued to be represented this way in the model.

In the case of injury, the necessity & usefulness of representing this data through specifically-modeled pain-based bodily feelings is particularly clear: it allows creatures to use their internal model to identify what part of their body requires immediate attention, and to help direct subsequent behavior accordingly. Again, as we've explained, the initial injury itself cannot direct that attention—it *must* subsequently generate the pain so there is a specific "marker" within the model to direct that attention. Instead of the injury generating, say, a visual response that causes the injured area to flash red within the model, the brain (for the reasons just explained) uses bodily-based sensations to direct that attention within the model.

In truth, the necessity & usefulness of representing this data through specifically-modeled pain-&-pleasure-based bodily feelings becomes particularly clear when you take a closer look at almost *any* of our feelings-generating mechanisms. For example, remember that *potential strangler* from Essay #5? Remember how their decision to strangle or not-to-strangle was necessarily *preceded* by the generation of the intent-defining *anger* that resulted from the observational thought "*He's trying to hurt my child!*" Here again, the subsumed model-based data that generates that observational thought cannot (alone) target an immediate responding angry action *within* the model. For all the same reasons described above, the resulting observational-thought-generated anger must be subsequently represented *within* that model (as bodily-based feelings) in order for both the anger & any visual focus on whomever is doing the hurting to be simultaneously subsumed and then help define the target of the subsequently-instructed strangling (and to help define the *intent* that spurs the choice to strangle).

Finally, the language-enabled addition of self-awareness in humans makes the presence of feelings within this model exponentially more useful—because we can specifically analyze those feelings that are generated by the experiences & dialogue that we process via our model, and sort through those sources & causes as we shape future choices & actions. As we discussed at length in our essay on emotions, *feelings* are not some ancillary component of conscious experience—they are powerful & necessary cognitive, decision-making tools, and thus, they are a powerful & necessary element of that experience-defining internal model.

The machine creates a modeled, emotion-soaked, self-narrated world (& body) and perceives it—and because we are that machine, we experience the perception of that modeled, emotion-soaked, self-narrated world & body. This is our consciousness. So we ask again, why *wouldn't* a machine like the human brain experience a perception of its own existence? How could the brain perceive & process a data format like bodily-based sensations *without* somehow *experiencing* those sensations within the model? And how could those sensations be of any use without simultaneously experiencing all of that *other* sensory data within the model? Ultimately, in a system such as this, if the machine *wasn't* capable of experiencing the data within this model, then it technically wouldn't be demonstrating any real capacity to *perceive* that modeled data, making the whole modeling process pointless—because in a mind with a modeled world, experience & perception necessarily go hand-in-hand. In the end, there is simply no compelling reason why they should not.

And there's one *very* compelling reason why they should: our own daily lives & ongoing existences provide mountains of proof that the perception of data in this model and our conscious experience are inextricably interwoven. A machine is as a machine does. And this machine does conscious experience via perception of a model, and we are that machine. That's why The Hard Problem doesn't really exist, and why we humans experience our existence.

###

A Post-Script Regarding Whether or Not Insects Actually Experience Their Existence

As we noted in the essay, Colin Klein & Andrew Barron's 2016 paper "Insects have the capacity for subjective experience" makes an impassioned argument *in favor* of the hypothesis that insects actually experience their existence. And their argument is strong enough to merit some discussion regarding why I disagree (although not vehemently) with their conclusions.

First, one of the reasons why I don't *vehemently* disagree is because their general view of what constitutes (and is required for) true conscious experience is essentially the same as Narrative Complexity's. They also deem that a central, integrated, multi-sensory spatially-based model of a creature's world & body is the primary necessary component of conscious experience. They just hypothesize that such mechanisms are similarly present in both humans *and* insects. Thus, even if their hypothesis is correct and ours (that insects don't possess subjective experience) is flawed—that still does nothing to undermine our explanation of why humans experience their existence. It merely means that we have underestimated the capacities of insect minds and in some future version of this essay will have to replace our insect example of a non-conscious mind with that reliably-familiar non-conscious old friend of ours: *C. Elegans* (whom we passed over in favor of bees in this essay because bees *appear* to provide a more interesting example of how complicated a "robot" can be without actually experiencing consciousness).

All of which means this post-script is making some fairly esoteric distinctions in the big scope of what this whole book is trying to define—but since we're in a post-script to an addendum, it's not a bad place to make some esoteric distinctions. The primary distinction at hand is that Klein & Barron's hypothesis is based on a theory by Bjorn Merker that identifies the midbrain's *colliculus* region (known to integrate visual & spatial data) as the center of conscious experience—the location that is most primary to generating a model of our world & ourselves. Narrative Complexity's hypothesis is based on Edelman's pre-frontal-cortex-focused Dynamic Core theory. In our view, the colliculus is clearly a necessary & central contributor to the broad multi-sensory network that generates that Dynamic Core-based model of conscious experience, we merely do not see it as *the central* component of that modeling network.

(And again, this is an essentially esoteric distinction within the big scope of our theory because even if Merker's theory is more accurate than Edeleman's, Narrative Complexity's own systematic explanation of these mechanisms would remain essentially unperturbed. We'd merely have to relocate our mechanism's elements to slightly different parts of that neuroanatomical network, but the order & fashion in which the data flows and is processed would be basically the same.)

However, this distinction is important to Klein & Barron's view of insects because they've identified a fundamental similarity between the general structure & functionality of the vertebrate colliculus and corresponding neural structures in insects. Thus, if Merker is correct, it might very well be possible that—because insects possess a similar "consciousness centerpiece"—bees & their brethren are actually experiencing their consciousness. And they make a strong case for tying together these systems in both insects & humans, and for these systems' vital roles in behavior & actions. Nonetheless, I do not believe they make a convincing argument that this spatially-focused arena must necessarily generate & subsume a full multi-sensory *model* of the environment (instead of just processing this partly-integrated but purely-visuospatial data in its "raw" non-modeled form) to play those vital roles in behavior & actions. And none of this even begins to speak to the issue of *feelings*—which are central experiential & functional components of vertebrate consciousness (& one of the main drivers necessitating an integrated multi-sensory model) yet would not appear to play any necessary role in Klein & Barron's version of insect consciousness.

In other words, even if bees (& other insects) are processing sophisticated visual & spatial data in similar locations & ways as vertebrates (and integrating that data with behavioral & action responses) their theory still doesn't demonstrate any true need to generate & subsume a unified model from this data in order direct insect cognition (instead of just handling the data in its raw form). In our view, all of those *bee-havioral* examples & mechanisms that we provided in this essay would still be coherently explained by our hypothesis within a system that's structured like Klein & Barron's but *does not* ultimately generate a model that is consciously experienced. From this perspective, the narrow-but-distinct chasm between Klein & Barron's view of insect consciousness and Narrative Complexity's view can be reflected in a fashion that's similar to (but more definitive than) the way we described the Free Will Paradox: Insects *seem to* have, for *most* intents and purposes, a subjective conscience experience—except that they technically (& *actually*) don't.

###

APPENDICES | *Distinct States of Consciousness & Non-Consciousness*

Rudimentary Map of Human Consciousness

Quick Sketch of Pre-Language Mammalian "Cognition"

Distinct States of Consciousness & Non-Consciousness

\ System State \	Brain Stem Systems	External Pain Input/ Amygdala Circuit	External Olfactory Input/Basal Ganglia Circuit	External Auditory & Tactile Input/ Thalamus Circuit	Thalamo-Cortical Loop & "Dynamic Core"	External Visual Input/ Thalamus Circuit
Stable Consciousness	Fully Active	Fully Active	Fully Active	Fully Active	Fully Active	Fully Active
Unstable Consciousness*	Fully Active	Partially Active	Partially Active	Partially Active	Partially Active	Partially Active
R.E.M. Para-Consciousness	Fully Active	Partially Active	Partially Active	Partially Active	Partially Active	Inactive *Suppressed*
Non-R.E.M. Para-Consciousness	Fully Active	Partially Active	Partially Active	Partially Active	Inactive *Suppressed*	Inactive *Suppressed*
Brain Shock Unconsciousness^	Fully Active	Partially Active	Partially Active	Inactive *Suppressed*	Inactive *Suppressed*	Inactive *Suppressed*
Reflexive Comatose Unconsciousness	Fully Active	Partially Active	Inactive *Non-Functional*	Inactive *Non-Functional*	Inactive *Non-Functional*	Inactive *Non-Functional*
Non-Reflexive Comatose Unconsciousness	Partially Active	Inactive *Non-Functional*	Inactive *Non-Functional*	Inactive *Non-Functional*	Inactive *Non-Functional*	Inactive *Non-Functional*
Catastrophic Unconsciousness	Inactive *Non-Functional*	Inactive *Non-Functional*	Inactive *Non-Functional*	Inactive *Non-Functional*	Inactive *Non-Functional*	Inactive *Non-Functional*

* Narrative Complexity defines **Unstable Consciousness** as state in which an individual alternates between conscious responsiveness (verbal and/or physical) and non-responsiveness—due to the effects of fatigue, sleep disorder (i.e., sleep-walking), injury, neural dysfunction or drugs. Unstable Consciousness is also highly vulnerable to slipping into some state of Para-Consciousness or Unconsciousness (depending upon the cause of the instability).

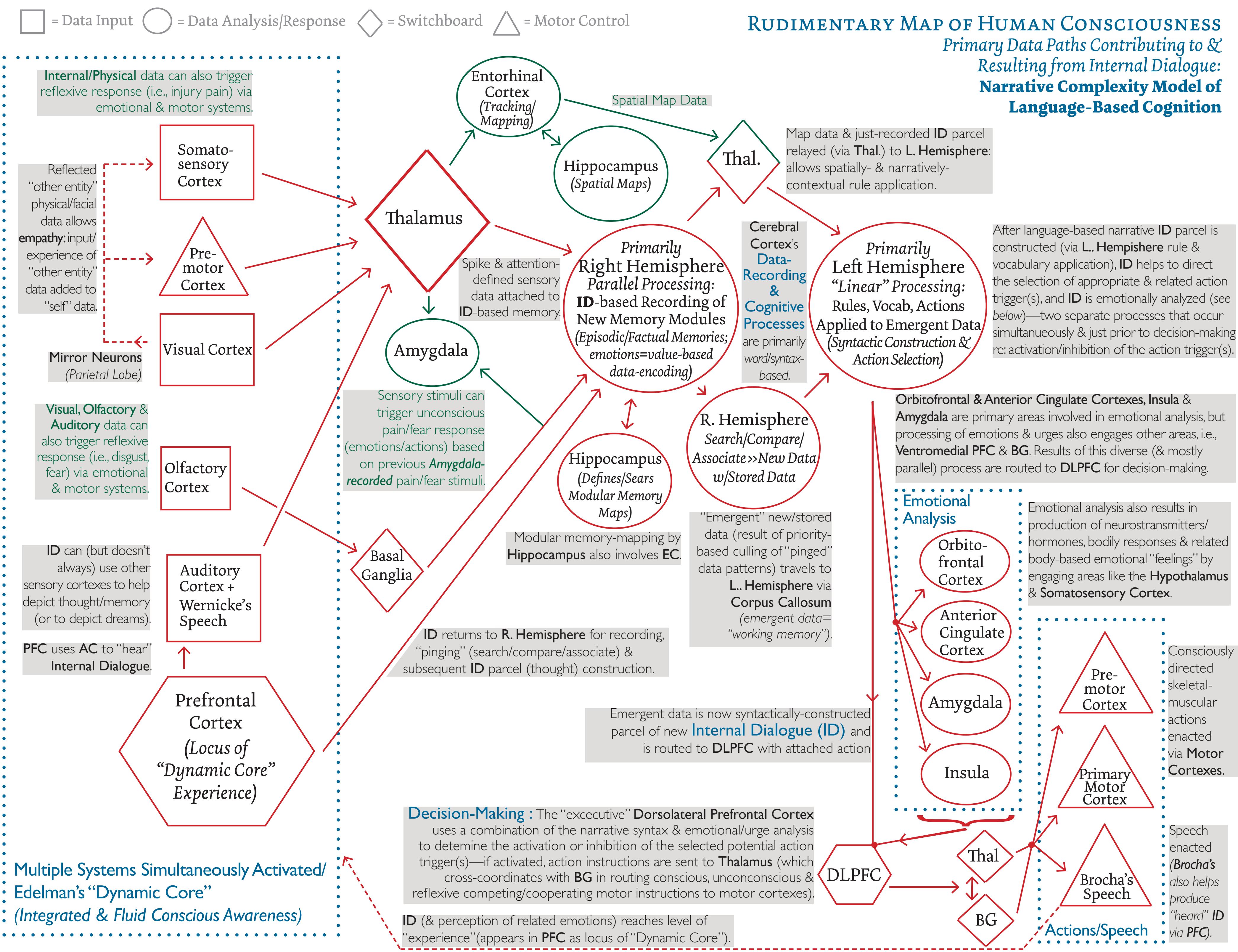
^ Narrative Complexity defines **Brain Shock Unconsciousness** as state in which the brain triggers an automatic (& temporary) shut-down or purposeful *suppression* of consciousness-generating & sensory-related thalamic activity in response to sudden & extreme concussive events, physical pain or emotional shock. We hypothesize that this state is *not* inactivity due to temporary trauma-induced non-function, but rather that is an *adapted* state whose purpose is to prevent seizures that might result from attempted thalamic activity during temporarily resource-deprived (likely oxygen-deficient) neural circumstances (caused by concussive events, physical pain or emotional shock). Anecdotal evidence suggests that this "self-initiated" & temporary thalamic shut-down state is unique to advanced primates & birds, and likely represents a recent evolutionary development. This state can be distinguished from Para-Consciousness by the fact that Brain Shock individuals can still sometimes be awakened by olfactory stimuli like smelling salts (basal ganglia-routed stimuli), but they cannot usually be awakened by auditory or non-pain tactile stimuli that can easily rouse someone from Para-Consciousness, like yelling or shaking (thalamus-routed stimuli).

Regarding Para-Conscious & Unconscious Motor Responses: Barring specific injury to the motor cortex or spinal column, individuals should still possess a capacity for reflexive motor actions in response to intense external stimuli, depending upon what specific stimuli-circuitry are still active & functional in that particular state of non-consciousness.

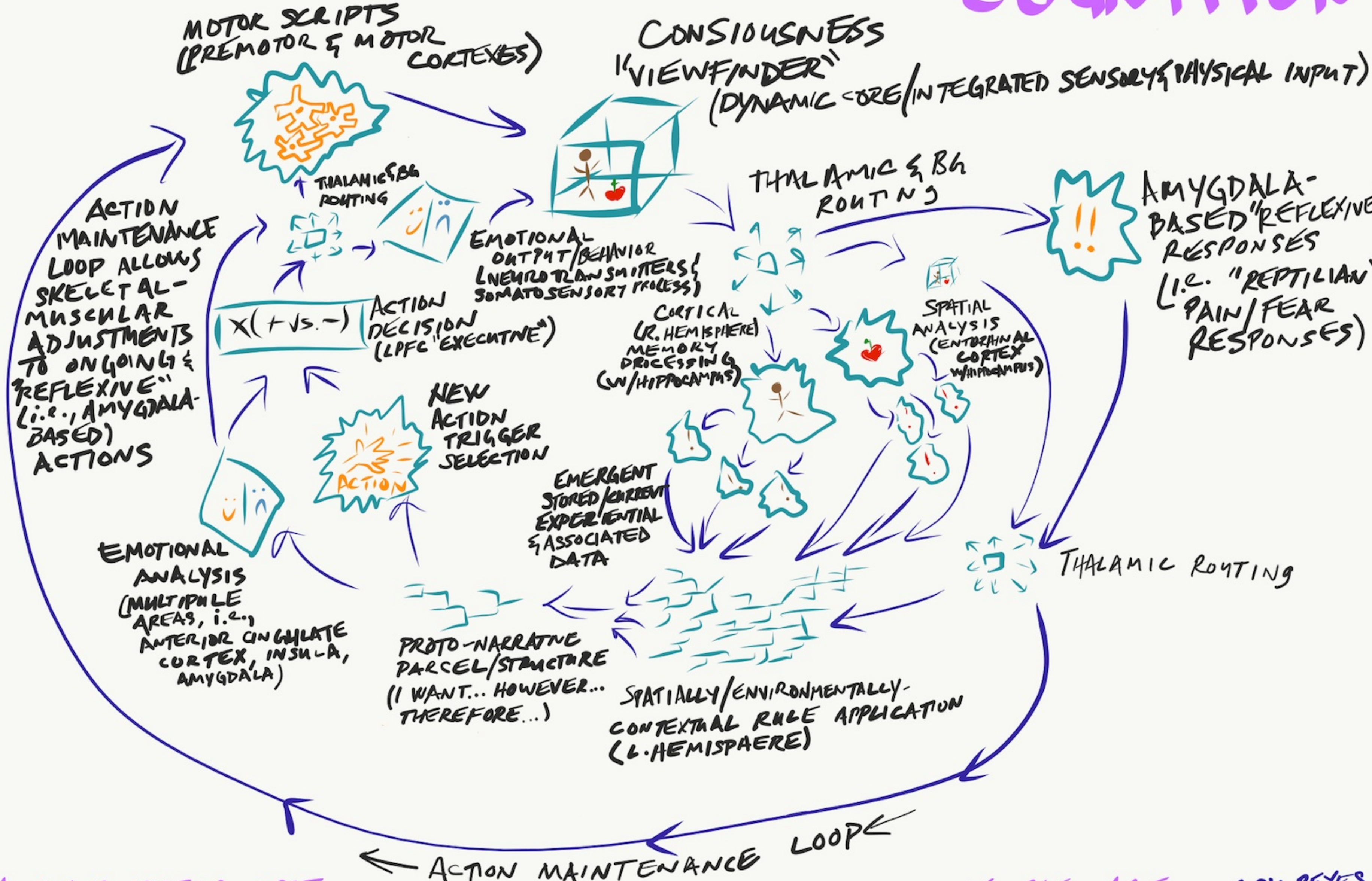
Rudimentary Map of Human Consciousness

Primary Data Paths Contributing to & Resulting from Internal Dialogue:

Narrative Complexity Model of Language-Based Cognition



PRE-LANGUAGE MAMMALIAN "COGNITION"



MAMMALS SPEND MOST OF THEIR TIME ACTING IN A WORLD THEY KNOW - INTERACTING WITH THINGS THEY ARE FAMILIAR WITH.

-R. SAL REYES
4/18/17