doi.org/10.53975/qilk-vgj4

Seeds of Science

Building a Brain: An Introduction to Narrative Complexity, a language & internal dialogue-based theory of human consciousness

R. Salvador Reyes

In the 2021 pre-print of his paper *"A voice without a mouth no more: The neurobiology of language, consciousness, and mental health,"* Jeremy Skipper from University College London proposes that "language and inner speech…cause and sustain self-awareness and meta-self-awareness, i.e., extended consciousness."[1] He goes on to layout "a 'HOLISTIC' model of the supporting neurobiology that involves a 'core' set of inner speech production regions that initiate conscious 'overhearing' of words."

In his conclusion, Skipper declares: "This review begs for a reconsideration of the role of language and inner speech in not only consciousness, but in cognition in general." Skipper's declaration is well-timed, because just such a "reconsideration" has been underway for the last decade—producing a new & comprehensive theory of consciousness that is built around a self-perpetuating loop of inner speech. (Because of the way this inner speech interacts with *itself,* I refer to it as *internal dialogue*.)

I propose that this *internal dialogue-based* looping model of consciousness & behavior, Narrative Complexity, provides a framework for effectively defining the specific (and central) role of language and inner speech within consciousness & cognition, and in turn, within the mechanisms of emotion & decision-making that consciousness & cognition help to govern. In addition, I suggest that Narrative Complexity's model supports (and is supported by) myriad recently-emerging & provocative consciousness-related models—ranging from theories like Skipper's Holistic model to Erik Hoel's dream-explaining Overfitted Brain Hypothesis[2] and Edmund Rolls' model of emotions and decision-making[3].

Laying out the full framework for such a theory & detailing its specific connections to those aforementioned models (and others) is more the domain of a book (and it is, "Narrative Complexity: *A Consciousness Hypothesis*"), but we can do the work here of introducing some of the theory's primary premises and outlining the model enough to demonstrate that it's novel, plausible, systematically comprehensive, and worth exploring in greater detail.

Building a Brain From the Beginning

Before we get started, however, I'm going to ask you to do something a little counterintuitive: *forget everything you know about consciousness*. Not because you don't know many possibly useful things about the many theories of consciousness, but because *there's just so much of it*. Introducing the framework for a new theory of consciousness by determining all its relationships to other major theories is a good way to make your head explode.

So let's not do that. Let's pretend you've never heard of this thing called consciousness. In that case, we'd need to start from the beginning—we'd need to build a brain from scratch and show exactly where this "consciousness" fits in, what it's doing there, and how language & internal dialogue enter the picture. To make things even simpler for your now-wiped-clean mind, we'll build this brain primarily according to its essential functions & capabilities, instead of focusing on its anatomy.

And if we're starting from scratch, a good place to begin is that delightfully reliable & ancient nematode, *Caenorhabditis elegans.* Old *C. elegans* (the original version—a kind of *alpha-C. elegans*—first appeared about 700 million years ago) and humans belong to slightly different evolutionary branches, but there are surprisingly common threads between our brains (e.g., *C. elegans* employs some of the exact same kinds of serotonin receptors as humans[4]). And this tiny worm's brain is a good example of the most fundamental functions and capabilities that neural systems allow: it uses external sensory organs to take in data from its current environment, and employs internal neurally-based processing to trigger (via tools like *serotonin*) an action-based behavioral response that is specifically geared toward their environment based on the data. In other words, when the worm's sensory input detects food, the worm's neural systems allow it to respond with specific behavior that aids it in consuming the food (like slowing its locomotion so it doesn't scoot past its meal).

Little *C. elegans* is also using simple internal data to shape that action/behavioral output. For example, if internal data indicates the worm is extra hungry (i.e., their systems are experiencing a greater deficit in resources) their locomotion is further slowed in order to spend more time consuming the food.[5]

Pretty basic stuff. So basic, in fact, that *C. elegans* has no conscious experience of any of this. That's partly because their neurally-based processing isn't really working as a *loop* yet. It's more like a *U-turn arrow*—taking in raw sensory data at the front end, using neural systems to process the data (with the help of related internal data), then launching situationally-appropriate behavior out the other end. Even as our chordate branch of evolution began producing invertebrate brains more complex than this simple nematode's, they still employed a similarly-structured (although more robust) set of mechanisms to achieve a similar kind of input & response behavior. Basically, although advancing invertebrate brains were capable of processing more information from their environment, and producing greater varieties of behavioral responses—their neural systems were still working like a *U-turn arrow* instead of a true *loop*. Which is, according

to our theory, partly why (just like *C. elegans*) our invertebrate ancestors never developed a conscious experience to call their own.

The Stuff of Conscious Experience: Adding a Loop & an Internal Model

Why does the absence of a loop translate into an absence of conscious experience? The answer takes us to our next step in building a brain: adding that loop. But before we do, we first need to add the loop's *raison d'être,* the other key component required for conscious experience: *an internal model*. (Requiring an internal model to achieve conscious experience is not unique to our theory; it's a key component of several other theories, including Bjorn Merker's [6] —who I mention for the sake of fellowship, because he is, like I am, an independent theorist without an academic affiliation.) During the epochs when those consciousness-deprived invertebrates were transforming into creatures with actual spines, something happened inside of their evolving brains—and evidence of it can be found in the earliest of vertebrates, those jawless pre-fish members of our chordate family tree: lampreys (or, to be precise, something happened inside the evolving brains of some original version of *alpha*-lampreys that appeared at the beginning of vertebrates).

These *alpha*-lamprey brains had a high-class problem: an abundance of useful incoming data about their environment. In particular, they were separately taking in an abundance of electro-sensory *and* visual data—both of which could be used to do the same thing: identify and track other creatures and objects within their watery surroundings. (Lampreys possess a lateral line electroreception system that takes in electrical data from their environment.) And it is here that the algorithms of evolution provided *alpha*-lampreys (and all the rest of us conscious creatures) with an extraordinary gift: the ability to integrate those multiple raw sensory data sources in the creation of a data-rich, three-dimensional *internal neural model* of that watery environment around them. Recent research on lamprey brains has revealed that they do this by integrating their electro-sensory data with visual data in the optic tectum via the dorsal thalamus & medial pallium[7] (the latter of which will contribute heavily to the development of the modern hippocampus—and, spoiler alert, the evolved versions of all these brain areas also play key roles in the human loop of consciousness).

Now our build-a-brain has a system that receives external sensory input and—*before it sends that data to be processed by those neural mechanisms that determine & trigger situationally-appropriate behavioral responses*—it *integrates* the data from those multiple sensory sources into a unified, dynamic & fluid internal model of that outside world. In addition, because this model is intended to track things in order to help direct how the creature physically interacts with them, the model also requires the integration of some of that *internal data* about the creature's body—essentially placing the creature itself *within* the world of their internal model.

For example, within this model the external data depicting another object in the creature's local space must be integrated with internal data depicting physical contact with that object in order for their brain to use this model to determine if they have touched something (and to determine what *specifically* in their environment they have

touched). This means that *U-turn arrow* has now necessarily transformed into a *loop*—outputting the products of its data-response processing not only as real-world action & behavioral results, but also as information (about its own body, actions & behavior) that is fed back into the internal model from which some of that data was originally derived.

According to Narrative Complexity's hypothesis, it is this rapidly-looping interchange and integration of external sensory & internal data—perpetually traveling from an internal model to response processing then back to the internal model, beginning the loop again—it is this process that constitutes and enables conscious experience. In short, as long as data appears & is integrated within that internal model, it is in some way a part of that creature's conscious awareness. This rudimentary version of consciousness isn't yet *self-aware* in the same way that humans' minds will be—but now that our build-a-brain possesses an internal model and a looping system, it at least has the tools to begin making that journey.

(At this point, if you've accidentally un-forgotten some of what you know about consciousness, you might be pondering questions about how all of this addresses David Chalmers' "Hard Problem"—although Chalmers' conundrum does not fall within our purview here, the full theory discusses why this looping system of perception enables experience.)

Expanding the Brain's Repertoire: Adding Memory & Emotion

S*elf-awareness* in our build-a-brain is still a few steps away, but at this point we *can* add another key mechanism that makes fantastic use of our newly-installed conscious experience: *memory.* By adding a memory mechanism our brain can now *record & store* some of that consciously-experienced data from their internal model. Similar to how those earlier brains helped to shape behavioral responses to incoming data by applying related internal data about their bodies (e.g., *C. elegans* slowing its roll when hungry), evolving vertebrate brains (specifically, early amniotes[8]) began to help shape behavioral responses to incoming data by checking that data against related & *internally stored* data that was recorded from previous high-impact (in neural terms, *spike data*) conscious experiences. Which is a very long sentence that basically says: *creatures could remember important stuff that happened, then apply it later when determining what to do in similar situations.*

This is an incredibly useful capability—and to really take advantage of this new memory mechanism, we need to add another key component to our build-a-brain: basic *emotions*. For the purposes of my theory, I like to refer to these early, pre-primate emotional mechanisms as *proto-emotions* (partly because, as the full theory explains in detail, complex modern human emotions all evolved out of these earlier simpler versions). As mammals quickly discovered, when you combine an internal model with memory and emotions in a looping system, consciousness begins to create some serious magic.

Emotions make the most out of memory by allowing a creature's brain to use *pleasure & pain* to, essentially, tag stored environmental & experiential data—other creatures, objects, and events—as *helpful or harmful* (*aka*, correlate experiential data with a *gain or loss* of some kind). This is what allowed memory data to be of use when checked against related current incoming data in the process of shaping behavioral responses—because memory data could help, for example, inhibit behavior or interactions that previously resulted in harm (*an experienced loss*) during similar circumstances, or vice versa.

As an added bonus—due to the way in which such a system is forced to work—these proto-emotions served to *enrich* conscious experience itself. That's because emotional data *must* somehow be represented within that consciously-experienced model in order to do its job (which has led to bodily-based "feelings" to become a primary element of the model). If emotional data is going to be correlated with an experience (*aka, a creature/object/event*) that's being depicted within that internal model, then the emotional data must *also* be output into the model (as"feelings") *after* the emotions have been produced by our data-response processing. This allows both the emotions & the correlating experience to be *simultaneously present* within the model—which ensures that the emotional data & the correlating experiential data can be *jointly* routed back out of the model and *associated* with each other during subsequent data recording & response processing.

Which is pretty complicated way of saying that if you kick a dog, the pain (& *anger*) caused by the kick will emerge in their model and be consciously experienced by the dog. And if the dog is also depicting you within its model while it is being kicked and experiencing that pain & anger, it's going to associate a recorded memory of your image (& scent) with those emotions. Which is a bad thing for you if you don't want that specific dog to bite you in future encounters. (It's also a bad thing for *this* encounter, because the dog's *anger* might spur it to bite you *right now*—revealing how emotions are designed to help shape both future *and* current behavior.)

Complicating Cognition: Adding Creative Responses & Learned Behavior

Now our build-a-brain is really humming. It can: depict & experience the surrounding environment and its own body within that environment; depict & experience different kinds of pain & pleasure in response to a variety of different gain & loss events (*proto-emotions*)*;* and record & store spike data from those experiences for future reference when using its data-response processing to determine situationally-appropriate behavior. In other words, compared to all the brains that came before them, evolving mammals had it pretty good.

In fact, now that all of those mechanisms are in place and working together, if we let evolution do its thing and increase the power & complexity of our data response-processing, we can add something very, very cool to our system: dynamic & creative behavioral responses and *learned* behavior. As that processing power & complexity increase, all of this varied recorded data involving a creature's experiences & actions begins to become more specific in its detail, allowing for more complex

categorizations, associations & divisions within that data. In other words, this data becomes more *modular*—i.e., the data strings are broken into more chunks of self-contained, independently-associative component parts.

What does that mean? It means that advancing mammals like dogs are able to independently associate a single event within a longer sequence with the specific result of the entire sequence—leading to an event response that correlates to the actual end result of the previously-experienced sequence and not the specific event (aka, rudimentary *symbolism*). *Pavlov's dog*: ring the bell and the dog salivates in response to the expected specific result of the sequence, the food—not because he wants to eat the bell.

This kind of modularly-managed, learned behavior also opens the door (in ever-advancing mammalian brains) to dynamic problem-solving and creative responses. This is because—by breaking longer sequences into their smaller component action & event chunks—these brains are able to combine specifically-useful component parts of different sequences to create novel new sequences in response to common challenges. If the newly-created response is successful, the positive emotions produced by the resulting gain can help to strengthen the recording of the sequence, enabling its future reuse in related circumstances—*aka,* learned *(& novel)* behavior.

The Arrival of Human Consciousness: Adding Language & Internal Dialogue

At its core, the overarching result of all this brain development can be boiled down to one main benefit: these kinds of advanced minds are much better at making accurate *predictions* about *what is going to happen* based on *what is happening right now* (and by applying any usefully-related memories regarding *what has happened before*). And when a creature's brain is capable of employing this kind of modularly-based predictive process in a wide variety of highly robust & diverse ways, we have a specific name for those creatures: *primates.* (And sometimes corvids, or cetaceans, or advanced cephalopods—but let's stick to our family tree here.)

Yes, our build-a-brain journey has reached the precipice of its final destination: the human mind. After the most advanced monkeys & great apes finally developed their rich capacity for generating a wide range of dynamic & creative behavior via their increasingly complex & modular data-response processing, hominins only needed to add one last piece to put this amazing system of cognition & consciousness over the top: *language & internal dialogue*.

Luckily, our closest ancestors inherited a brain that was ripe for the introduction of language & internal dialogue. All those modular predictions & dynamic responses in earlier primates had resulted in an important linguistically-friendly development: a kind of *proto-narrative syntax*. When a monkey identifies a desired piece of very ripe fruit hanging from a tree branch, but the branch is too flimsy to traverse and the fruit is too far to reach from the trunk or the ground, the monkey might respond to this conundrum by shaking the branch hard enough to dislodge the fruit, then retrieve it from the ground. This process, at its heart, is employing a kind of "if, then, however, therefore"

proto-narrative syntax to help creatively predict & execute what must be done in order to achieve the desired goal*. If I want the fruit, then I need to pick the fruit, however I cannot reach it, therefore I need to dislodge it… If I want to dislodge the fruit, then I need to shake the branch…*

All of that can be done without language, but such a proto-narrative syntactic process has its creative limitations—limitations that are *erased* by the addition of language & internal dialogue. So let's do it, let's finally add language & internal dialogue to our build-a-brain and see what we get. That proto-narrative syntax is already set-up to determine (assumed) causal relationships, to categorize & employ elements like "predicates" (*aka*, component parts of action sequences) and "objects," and to rearrange these elements in novel ways to achieve specifically-predicted results. These capacities are almost all we need to install language & internal dialogue in our system.

The only *other* required capacity is the development of vocal abilities that are sophisticated enough to generate a dizzying array of different highly-specific kinds of sounds. (*Fortunately, according to theories from researchers like Terrence Deacon, hominins had already begun developing those capacities thanks to other non-linguistic evolutionary forces.*)[9] With these rudimentary syntactic capabilities & increasingly-sophisticated vocal capacities in place, hominins were able to begin associating these basic syntactic components (like actions & objects) with specific, consistent vocalizations—aka, *words.* The seeds of language-based cognition had been planted. (My theory's preferred depiction of complex language-based cognition is the model presented in M.A.K. Halliday's & Christian M.I.M. Matthiessen's book, *Construing Experience Through Meaning: A Language-Based Approach to Cognition*.)[10]

Using these words to express *out loud* the components of that previously-silent syntax provided huge advantages in communicating & cooperating with other hominins (allowing for more specificity than earlier, simpler forms of inter-hominin communication like pantomime). In addition, the expression of these words *internally* provided huge advantages within our systems of *cognition & consciousness*. This is mostly because of our brain's very old pair of friends: the loop & the internal model.

Thanks to the kinds of mechanisms outlined by Skipper in his HOLISTIC model, hominins are capable of engaging this looping, associative, self-generating linguistic process *without* speaking the words out loud, but merely by internally "overhearing" the expression of those words *within their consciously-experienced model*. And after the internal expression of an unspoken (but still linguistically-based) thought is integrated into our model, that thought is routed back into the loop and processed by those memory & data-response systems in order to help shape the loop's *next* round of emotional, behavioral & linguistic output (which will itself then be integrated into the model and subsequently subsumed—*round-&-round we go*). This looping process allows our system of consciousness to use *experience* (our internal model) as a conduit that receives a language-based thought from our generative cognitive processes and sends it to the memory systems that will subsequently record & associate that thought

—thus allowing its data to seed & inform the *next* generation of thought produced by the loop.

And this experience of "overhearing" (essentially, internally perceiving) expressed-but-unspoken language can actually occur via any sensory avenue present within the model—e.g., the visual or physical/tactile avenues employed by individuals with hearing & other deficits. Which reinforces the conclusion that it is *the language itself* (employed however possible) and its presence (in whatever form) within our consciously-experienced internal model that allows consciousness to leap from merely *experiencing unawares* (pre-language mammals) to *experiencing the experience of being aware of your experience* (humans). The extraordinary significance of that leap might best be expressed by someone who actually *remembered* making such a leap themselves: Helen Keller, who was born blind & deaf and didn't acquire language until middle childhood. In Skipper's paper, he presents a telling quote from one of Keller's own books describing her experience:

> *"Before my teacher came to me, I did not know that I am. I lived in a world that was no world. I cannot hope to describe adequately that unconscious, yet conscious time of nothingness. I did not know that I knew aught or that I lived or acted or desired. ... Since I had no power of thought, I did not compare one mental state with another. … When I learned the meaning of 'I' and 'me' and found that I was something, I began to think. Then consciousness first existed for me."* ("The project Gutenberg eBook of the world I live in, by Helen Keller," n.d.)

Keller's experience mirrors the experience described in Susan Schaller's book, *A Man Without Words*[11]*,* which tells the story of a deaf man who finally learned word-based sign language after living with a group of deaf individuals who only communicated via basic, communally-shared & -developed pantomimes. After discovering words, the man could barely describe the "dark time" of his pre-language life. In light of such descriptions, we can confidently declare that—by adding language & internal dialogue to its self-experienced model—our build-a-brain's consciousness has finally developed that magical glow: *self-awareness*. Behold the human mind: an entity that *experiences the experience of being aware of its experience.* And, as Keller attested to, a mind cannot have that experience if it doesn't include *some kind* of self-generated & self-perceived, language-based thought.

Yes, earlier creatures had long been consciously experiencing all those other elements (sensory input, internal physical data, emotions) that were integrated into their internal model, but their self-generated cognitive responses (spurred by all that input) resulted primarily in *actions.* The cognitive code that determined the actions (like our monkey's proto-narrative syntax) was not in any kind of data format that could be integrated into the experiential model, and therefore these cognitive "thoughts" that helped to shape action remained unexperienced. Cognition in language-less mammals could produce action & emotion, but the cognitive code that generated those results remained silent.

With the introduction of language & internal dialogue, that cognitive code (literally) found a voice. Hearing & experiencing within our minds the code that shapes our actions and behavior has had a transformative impact on humans, sent us on a journey like no other. Humans think *to themselves* about their world & its challenges using language. The experiences of all those emotions, actions, interactions, environments, objects & predictions are *expressed* (and intricately recorded & associated) using a sensorially-experienced, highly-complex, word-based neural code. And humans *know* that they are thinking about things—*we are aware of ourselves & our thoughts.* We're telling ourselves the story of our lives.

But the transformative effects of language & internal dialogue aren't just limited to how we experience consciousness, they extend to how the system itself works—its efficiency and its ability to generate creative solutions. In terms of efficiency, consider this: the same words that cognition uses to shape action are then used by our model to shape experience, and then used by our recording/associating mechanisms to shape memory. This means the entire loop of conscious experience and unconscious thought-processing cognition can be tethered to a single kind of code: language. Memory, beliefs, emotion, association, calculation, imagination, decision, action—all of it can be enacted and managed using a narratively-constructed, word-based code that is being employed or manipulated in some fashion at every juncture of the loop. The most obvious follow-up question to this declaration is: *how, exactly?*

The full answer requires laying out in detail the myriad mechanisms that compose our loop—but to arrive at a much shorter answer, I first ask you to consider: *my favorite bird was stolen from my bedroom by a stranger in the middle of the night*. There's a lot of emotional content in that tiny tale: sadness, anger, fear. And I'm betting most of you automatically felt or perceived or inferred at least a tiny bit of some (or all) of those emotions when you read it—just by processing the words. That's your brain using the code of words (and their generative, semantic & associative powers) to calculate and produce emotion. The narrative structure and the words allow us to identify losses, the value of the losses, and predictions of future potential losses in like circumstances—these judgements and the feelings they produce are all the domain of emotions, and they can all be calculated using words. We take it for granted, but words are powerful stuff.

Words are also creative stuff. Each one is a modular unit defined by its own specific syntactic & semantic properties (function & meaning), and each experience of a specific word is capable of generating its own unique associations with other words and experiences when subsumed (as part of a thought) into our memory's recording/associating/comparing mechanisms. This means that words (& the phrases they build) have the capacity & opportunity to generate a multiplicity of connections & relationships between each other and the longer strings of patterns they create. All of this makes those robustly-attributed & specifically-defined—but still malleable & modularly-employed—words fantastic tools for creative tasks like cross-associating ideas, generating analogies, and deciphering metaphors. Essentially, words & language

are great for comparing patterns to identify uniquely useful cross-applications of novel patterns in different circumstances.
If you're looking for a creative solution to a difficult problem, human language and its word-based code have proven to be exemplary tools for such tasks. Maybe the best example: *math*. All of those brilliant equations explaining all those complicated (but highly useful) solutions are really just very complex predictive narratives (albeit ones that partly follow their own specific syntax & semantics in their construction). The language of math is simply another word-based way to express predictive causal relationships. *2 + 2 = 4* is ultimately a causal narrative that's not unlike *Jill pushes Jack and Jack falls down*. Again, we take it for granted, but words are powerful stuff.

Indeed words are so powerful that, in one fell swoop, language fundamentally changed the very nature of consciousness & cognition along an epic march of chordates that stretched back nearly 700 million years. With the addition of internal dialogue, that ancient loop of consciousness was able to make its own use of those powerfully malleable, symbolic tools, *words*—by internally generating, experiencing & subsuming observational & predictive, narratively-based sequences of language, and employing their component parts to create an array of different kinds of associations between the data and to identify broader overarching patterns among the sequences, and then using all that data to immediately create useful new (& often *introspective*) sequences via the same process. This is Douglas Hofstadter's *strange loop*[12]: a snake in a circle eating its tail, a perpetual motion machine that both generates and consumes itself via the loop—using that constantly-looping system to produce new thoughts from the previous thought's data, keeping the engine of language circling our minds and sustaining our self-aware experience throughout all our conscious hours.

With the addition of language & internal dialogue, the human mind's capacity to understand itself and its world—its *experience* of its consciousness—took an exponential leap. It's a leap that has made our mind capable of deciphering & understanding how its *own self* is capable of that very same deciphering & understanding—which we might consider the ultimate act of *self-awareness,* and may be the crowning *self-achievement* of this thing called *consciousness*.

What Can We Do With Knowing This?
That's a very good question. In the end, the truth is that understanding the central role of language in consciousness & cognition is just the beginning. But, with that understanding in hand, we *can* build a full model of a looping, language-&-internal-dialogue-based system of consciousness that lays out the workings of loop's consciousness-generating subsystems (like emotion, memory, cognition & decision-making) while explaining how all of these systems impact & define behavior & experience. (Which is why the full theory begins by continuing the discussion of language & internal dialogue that we've begun here, and then builds the rest of the model on that foundation.) And after we've assembled all of those subsystems into a comprehensive model, we can *apply* that model by using it to do things like specifically defining & identifying the roots of different behavioral disorders & mental illnesses, explaining the specific ways in which our belief systems impact our choices, or

determining the true nature of free will. (If we wanted to, we could also use the model to help build more "human-like" artificial intelligence, but I'm not sure that we'd want to, so I haven't done that yet.) These are the things that we can do when we begin with an understanding of language's looping role in consciousness & cognition—the knowledge is like a Rosetta Stone that can be used to decipher how the entire system is talking to itself. That's what *Narrative Complexity* does: it uses the Rosetta Stone of language (human cognition's primary code) to decipher the subsystems that generate consciousness & behavior, then builds a comprehensive model of the full system, and applies that model by using it to reshape our understanding of "real-world" issues like specific behavioral disorders & mental illnesses. So, although this piece *(& our build-a-brain)* ends here, this is really where the journey into consciousness *(& behavior)* begins…

~

You can explore visual depictions of Narrative Complexity's internal dialogue-based, looping model of consciousness in different levels of detail in the attached *Diagram of Narrative Complexity* and *Rudimentary Map of Human Consciousness (found after gardener comments and endnotes)*.

~

Author's Note: As indicated by the title, this is only an introduction to my full theory of consciousness & behavior. You have a lot more questions—I have a lot more hypothetical answers: "Narrative Complexity: *A Consciousness Hypothesis.*" [https://rsalvador.com/NC_Hypothesis_RSReyes.pdf]

Gardener Comments

Michael M. Kazanjian:
The article is a good summary of consciousness, and its relevant literature. Particularly welcome are the graphics. Visuals are all too often missing in most papers. in this case, they offer the optic opportunity for seeing what the verbal presentations wishes to convey. The chart in this paper enables readers to spatially follow the thought pattern in its complexity.

Dan James:
An ambitious paper that read more as a poetic speculation for me (though it lacked poetic brevity). Whilst the paper offered some citations, these did not convince me that it was accurately set in context, still less that those citations that were mentioned reflected the huge amount of contemporary research in this area. As a result, for me, a persuasive and coherent argument was not developed. The obvious self-promotion in the paper was also not helpful or good practice and I feel should have been omitted.

DK:
This is a very interesting article, I enjoyed reading it. I think that you can cite in the article "The bicameral mind" by Julian Jaynes, which is the first book that proposes ideas that are similar to what you discuss.

Anonymous:
I'm not an expert on consciousness and yet did not find much new in this paper. The writing is engaging, but juggles with terminology that is insufficiently spelled out. It reads as an advertisement of the book, but it merely tells what the theory (supposedly) can do, instead of actually showing it. In other words, there is insufficient here to judge the merit of the theory.

Kasra:
Absolutely fascinating piece! Yes from me: the author's synthesis of various lines of research seems novel, the writing is superb, and the author provides concrete examples of how Narrative Complexity helps advance not only our understanding of the brain but also (eventually) our technologies for healing/improving it.

Wrote a few comments below, they're mostly out of curiosity / to clarify my understanding.
--
Author: "[lampreys] were separately taking in an abundance of electro-sensory and visual data—both of which could be used to do the same thing: identify and track other creatures and objects within their watery surroundings. And it is here that the algorithms of evolution provided alpha-lampreys (and all the rest of us conscious creatures) with an extraordinary gift: the ability to integrate those multiple raw sensory data sources in the creation of a data-rich, three-dimensional __internal neural model__ of that watery environment around them." (3)

It seems like the crux of the matter (the crucial requirement for consciousness) is not the integration of multiple kinds of stimuli, but specifically the formation of an internal model. e.g. even in nematodes, there is integration of "different kinds" of information—homeostatic and sensory—but they do not have a 3-dimensional model, hence no consciousness. and on the other hand, couldn't there be (in principle) some organism that forms a 3d model of the world, but only does that with one kind of sensory input (e.g. olfaction)? so basically, I'm curious if the requirement is just the internal model, or if it's multisensory input AND an internal model. (or is it perhaps that only when there were multiple kinds of sensory input did the *evolutionary imperative* for an internal model appear?)

Author response to Kasra: Great question. I believe that conscious experience is enabled by an internal model that integrates multiple sensory sources—thus, neither capacity by itself would enable consciousness. This is because, according to my theory, an internal model that integrated multiple sensory sources would've required a novel

new (conscious-enabling) system for *perceiving* the integrated data that was presented in this newly unified fashion.

Previous to this integrated model raw sensory data was fed directly to response processing systems in order to generate behavior, because there was no need for an intermediary step—those systems had evolved specifically to process that incoming sensory data. But an internal model that integrates multiple sensory sources is generating a kind of data configuration that is very different from that raw incoming sensory data, thus requiring a novel new perceptual system that is capable of subsuming specific data in this model and sending it on to those response systems in forms that they can use to generate responses. My theory proposes that it's this intermediary step in the process—the perception of modeled & integrated sensory data—that gives rise to conscious experience.

In terms of the evolutionary necessities that led to this modeling, it probably had something to do with the fact that creatures like lampreys were using using more than one sensory mode to accomplish very similar tasks (e.g., visual & electrosensory to detect & track objects), making the integration of those sensory feeds more likely/useful/adaptive, and this integration led to 3D modeling because that was the most useful way to depict the integration of those two types of data for those purposes. Because this model was a primary part of the system & connected to myriad functions, integrating other sensory modalities became possible as creatures evolved. (A deeper discussion regarding the processing of raw sensory data vs. the processing of modeled data, and their relationships to conscious experience can be found on pgs. 246–252 of the full theory.)
--

Author: "Emotions make the most out of memory by allowing a creature's brain to use pleasure & pain to, essentially, tag stored environmental & experiential data—other creatures, objects, and events—as helpful or harmful (aka, correlate experiential data with a gain or loss of some kind)." (5)

Why is emotion necessary for this? Why not just tag the events with plain memories of pleasure and pain? Also, does this mean non-emotive creatures (like fish?) are incapable of producing any associative memories with stimuli? (the author does seem to clarify this later on when they say the brain can "depict & experience different kinds of pain & pleasure in response to a variety of different gain & loss events (proto-emotions)", i.e. while pain and pleasure are effectively one-dimensional, emotions allow you to add multiple new dimensions of valence?)

Author response to Kasra: Basically, emotions have two core purposes according to my theory: imprinting & signaling. Imprinting involves things like how strongly a memory is imprinted (based on factors like emotional intensity, or the level of basic pain or pleasure) and tagging the degree of loss/gain (also reflected by emotional intensity) when recording that data. For these purposes, that core loss/gain correlated

pain/pleasure element (which is at the core of every emotion) is the source of their functionality.

Signaling involves more specific kinds of effects—such as the way that different emotions encourage specific kinds of behavior: embarrassment encourages hiding from people, anger encourages lashing out at people, affection encourages physical (& other kinds of) closeness with people, etc., etc. Those emotions all generate some kind of pain or pleasure at their core, but the various aspects of the “feelings” that compose the specific pain or pleasure of these different emotions (feelings that are all bodily-based) are all slightly different, and those differences correlate to the behavioral differences generated by the emotions.

This gives the brain a wide array of different ways to respond to different kinds of circumstances, ways that are all calibrated to the specific emotional (gain/loss) variables that are present within those circumstances. (The fundamentals of human emotional mechanics are laid out on pgs. 23–48 of the full theory.)
--

The author also makes the point that rudimentary consciousness begins with vertebrates, but the referenced Merker paper points to recent evidence that some invertebrates (insects) *do* seem to form a multisensory internal model of their body in space, and hence they might actually be conscious? this is tangential to the main thrust of the paper, but curious for the author's opinion on this.

Author Response to Kasra: Indeed, although I share Merker’s belief that an internal model is the key to conscious experience, Merker’s theory places the superior colliculus at the center of that consciousness-generating process—while my theory aligns itself with Gerald Edelman’s Dynamic Core Hypothesis, which considers the prefrontal cortex to be the locus of this process. In many ways, this is an ultimately esoteric distinction (my model doesn’t argue against the colliculus playing a key role in this process, I simply wouldn’t declare that it plays *the* key role). Nonetheless, this does have implications in areas such as determining whether or not insects are conscious. (There’s not quite enough space here to go over those particular implications, but they are discussed specifically on pgs. 256–257 of the full theory.)
--

Also entirely out of curiosity: the author mentions experiences like those of Hellen Keller in which consciousness prior to language was a kind of "timeless, nonexplicit, nondifferentiated" awareness. It seems that peak meditative states (which often require prolonged bouts of concentration and the explicit attempt to "let go" of thinking) could be described similarly. Does the author think these are fundamentally the same kind of experience?

Author response to Kasra: I would say that this is *almost* true, but not really. That’s because, according to my theory, once we learn language we can never truly

experience any state of consciousness *without* some kind of language being involved at some level in the process. I understand that seems kind of hard to believe, but thinking is such a deeply intuitive process that it's very easy to not appreciate the fact that we're doing it *all the time*. Meditation is certainly a different kind of experience that involves a different manner of thinking—slower, more fluid, quieter, etc.—but it isn't truly language-less in the way Keller's experience would have been. In addition, Keller's description of being languageless centered around a formless feeling of *not even knowing* that she was an "I"—which is a kind of self-knowledge that you can't really unlearn once you know it, even via meditation. (Our never-ending inner thoughts & meditation are discussed further on pgs. 109–112 of the full theory.)

Endnotes:

1. Skipper, Jeremy I. "A Voice Without a Mouth No More: The Neurobiology of Language, Consciousness, and Mental Health." PsyArXiv, 16 Aug. 2021. Web.
2. Hoel, Erik. "The overfitted brain: Dreams evolved to assist generalization." Patterns 2.5 (2021): 100244
3. Rolls, Edmund T. Emotion and decision making explained. Oxford University Press, 2014.
4. David E. Nichols, Charles D. Nichols. "Serotonin Receptors." Chem. Rev. 2008, 108, 1614–1641.
5. Daniel Togo Omura, "C. elegans integrates food, stress, and hunger signals to coordinate motor activity." Massachusetts Institute of Technology, June 2008.
6. Merker, Bjorn H. "Insects join the consciousness fray." Animal Sentience 1.9 (2016): 4.
7. Pombal, M. A. "Afferent connections of the optic tectum in lampreys: an experimental study." Brain, behavior and evolution 69.1 (2006): 37-68.
8. Laberge, Frédéric, et al. "Evolution of the amygdala: new insights from studies in amphibians." Brain, Behavior and Evolution 67.4 (2006): 177-187.
9. Deacon, Terrence. *The Symbolic Species*. Norton, 1999.
10. Halliday, M.A.K., and Matthiessen, Christian M.I.M. *Construing Experience Through Meaning: A Language-Based Approach to Cognition.* Continuum, 1999.
11. Schaller, Susan. *A Man Without Words*. University of California Press, 1995.
12. Hofstadter, Douglas R. *I am a strange loop.* Basic books, 2007.

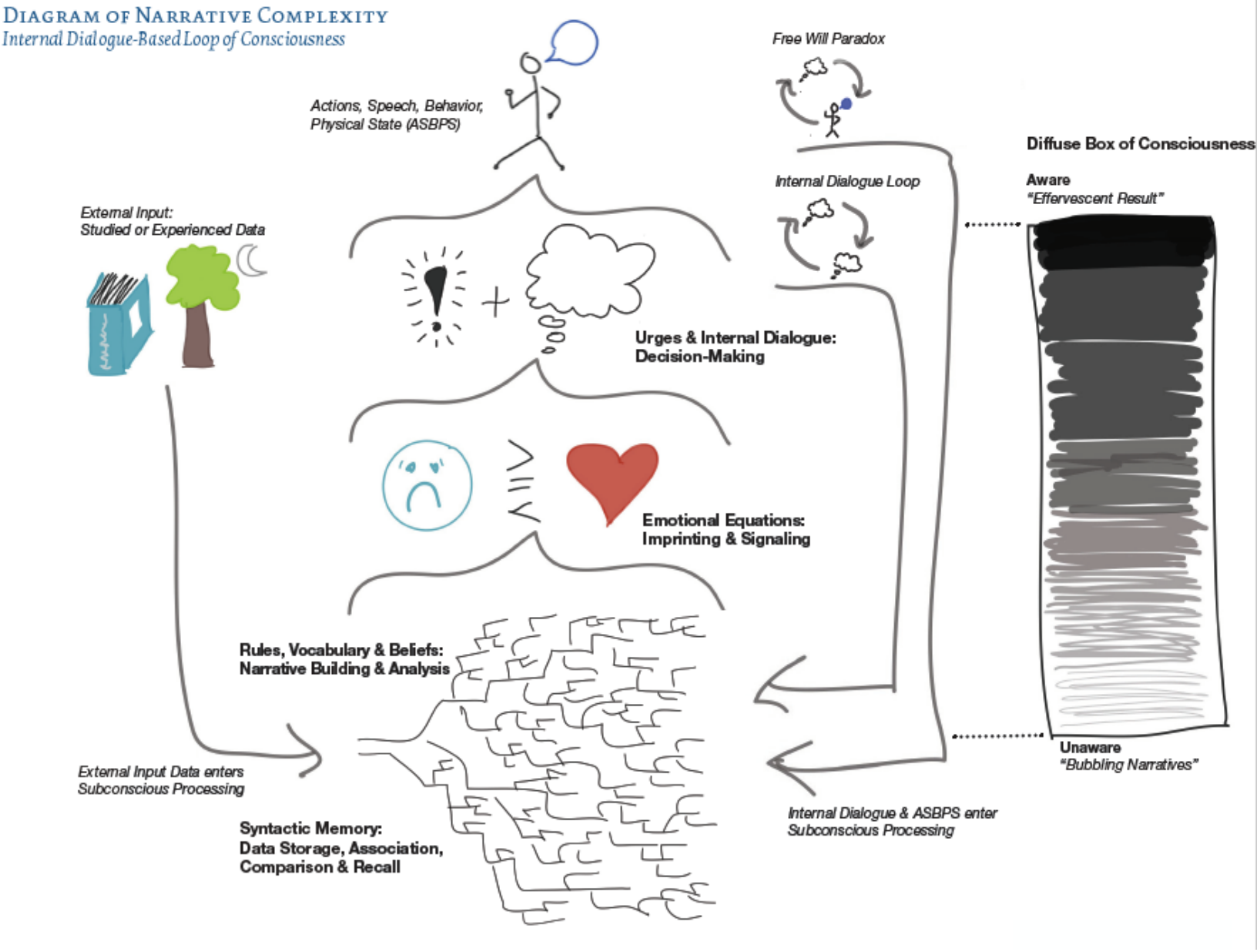
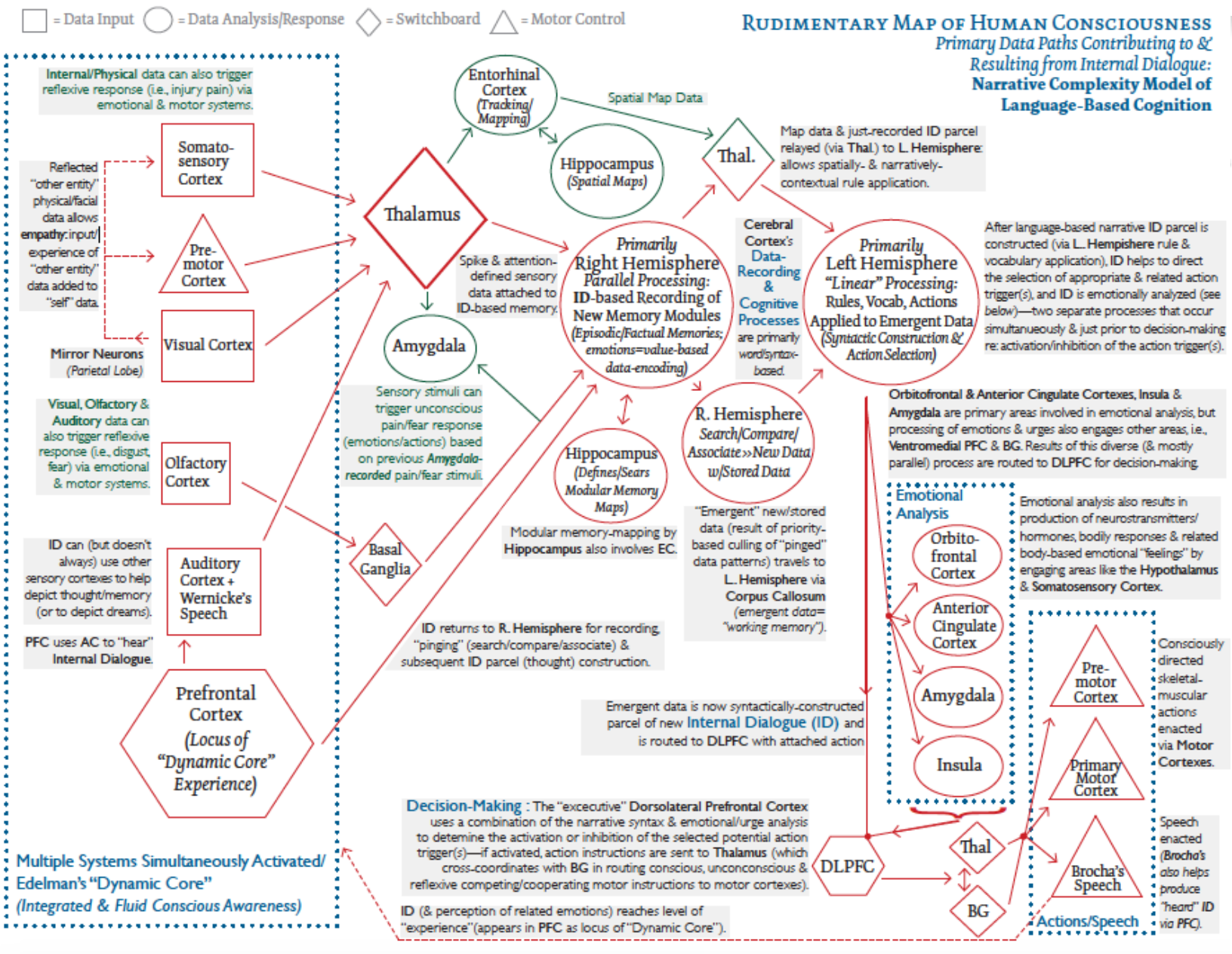


Diagram of Narrative Complexity
Internal Dialogue-Based Loop of Consciousness
Actions, Speech, Behavior, Physical State (ASBPS)
Free Will Paradox
Diffuse Box of Consciousness
Aware
"Effervescent Result"
Internal Dialogue Loop
External Input:
Studied or Experienced Data
Urges & Internal Dialogue:
Decision-Making
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


□ = Data Input ○ = Data Analysis/Response ◇ = Switchboard △ = Motor Control
RUDIMENTARY MAP OF HUMAN CONSCIOUSNESS
Primary Data Paths Contributing to & Resulting from Internal Dialogue:
Narrative Complexity Model of Language-Based Cognition
Internal/Physical data can also trigger reflexive response (i.e., injury pain) via emotional & motor systems.
Somato-sensory Cortex
Reflected "other entity" physical/facial data allows empathy: input/experience of "other entity" data added to "self" data.
Pre-motor Cortex
Visual Cortex
Mirror Neurons (Parietal Lobe)
Visual, Olfactory & Auditory data can also trigger reflexive response (i.e., disgust, fear) via emotional & motor systems.
Olfactory Cortex
ID can (but doesn't always) use other sensory cortexes to help depict thought/memory (or to depict dreams).
Auditory Cortex + Wernicke's Speech
PFC uses AC to "hear" Internal Dialogue.
Prefrontal Cortex (Locus of "Dynamic Core" Experience)
Multiple Systems Simultaneously Activated/ Edelman's "Dynamic Core" (Integrated & Fluid Conscious Awareness)
Thalamus
Entorhinal Cortex (Tracking/Mapping)
Hippocampus (Spatial Maps)
Spatial Map Data
Thal.
Map data & just-recorded ID parcel relayed (via Thal.) to L. Hemisphere: allows spatially- & narratively-contextual rule application.
Spike & attention-defined sensory data attached to ID-based memory.
Amygdala
Sensory stimuli can trigger unconscious pain/fear response (emotions/actions) based on previous Amygdala-recorded pain/fear stimuli.
Basal Ganglia
Primarily Right Hemisphere Parallel Processing: ID-based Recording of New Memory Modules (Episodic/Factual Memories; emotions=value-based data-encoding)
Cerebral Cortex's Data-Recording & Cognitive Processes are primarily word/syntax-based.
Primarily Left Hemisphere "Linear" Processing: Rules, Vocab, Actions Applied to Emergent Data (Syntactic Construction & Action Selection)
After language-based narrative ID parcel is constructed (via L. Hemisphere rule & vocabulary application), ID helps to direct the selection of appropriate & related action trigger(s), and ID is emotionally analyzed (see below)—two separate processes that occur simultaneously & just prior to decision-making re: activation/inhibition of the action trigger(s).
Hippocampus (Defines/Sears Modular Memory Maps)
Modular memory-mapping by Hippocampus also involves EC.
R. Hemisphere Search/Compare/Associate >> New Data w/Stored Data
"Emergent" new/stored data (result of priority-based culling of "pinged" data patterns) travels to L. Hemisphere via Corpus Callosum (emergent data= "working memory").
Orbitofrontal & Anterior Cingulate Cortexes, Insula & Amygdala are primary areas involved in emotional analysis, but processing of emotions & urges also engages other areas, i.e., Ventromedial PFC & BG. Results of this diverse (& mostly parallel) process are routed to DLPFC for decision-making.
Emotional Analysis
Orbito-frontal Cortex
Anterior Cingulate Cortex
Amygdala
Insula
Emotional analysis also results in production of neurostransmitters/hormones, bodily responses & related body-based emotional "feelings" by engaging areas like the Hypothalamus & Somatosensory Cortex.
ID returns to R. Hemisphere for recording "pinging" (search/compare/associate) & subsequent ID parcel (thought) construction.
Emergent data is now syntactically-constructed parcel of new Internal Dialogue (ID) and is routed to DLPFC with attached action
Decision-Making: The "excecutive" Dorsolateral Prefrontal Cortex uses a combination of the narrative syntax & emotional/urge analysis to determine the activation or inhibition of the selected potential action trigger(s)—if activated, action instructions are sent to Thalamus (which cross-coordinates with BG in routing conscious, unconconscious & reflexive competing/cooperating motor instructions to motor cortexes).
DLPFC
Thal
BG
Pre-motor Cortex
Primary Motor Cortex
Brocha's Speech
Actions/Speech
Consciously directed skeletal-muscular actions enacted via Motor Cortexes.
Speech enacted (Brocha's also helps produce "heard" ID via PFC).
ID (& perception of related emotions) reaches level of "experience" (appears in PFC as locus of "Dynamic Core").