Complexity Science Strategies for Conceptual Engineering

Brian Kobylarz

Abstract: The chief aim of this paper is an attempt to establish a conceptual engineering project where 'engineering' is taken literally. This demands a mass synthesis of discoveries made in cognitive science (conceptual processing, various theories of concepts) to provide a sense of the materials (what concepts are) along with engineering techniques found in complexity science (e.g. hierarchical modeling of concepts, redundancy, and robustness analysis). This paper is thus demonstrative, and each section builds on the previous. I first show how concepts are complex systems. A model of what I call concept-systems is then provided, and finally, I apply this machinery in a hypothetical conceptual engineering scenario, demonstrating that complexity can be useful to a conceptual engineer. This project pivots on helpful critiques Herman Cappelen has made, which any conceptual engineer must wrestle with. The idea is that by accepting Cappelen's challenges, there is a path through cognitive science and complexity science that can deliver us a possible conceptual engineering project.

Keywords: Conceptual Engineering; Concepts; Cognitive Science; Complexity Science; System Science

Introduction

Concepts perform essential jobs. These jobs are said to be involved with determining what we know, believe, perceive, categorize, conclude, infer, assert, detect, understand, love, hate, formalize, feel, and advocate. In Kantian terms, we may say that concepts are some of the preconditions of agency. Another way of putting it is that concepts precede results; they produce results in the most general sense—be they ordinary or extraordinary. Concepts are thus orientational, and we are as capable as our concepts. Examples can range from new pronouns for gender exploration to concepts that are ladened with political associations such as "criminal," "victim," or "refugee." Abstract examples exist too, such as those found in the sciences: consider "rational actor" and how entering an economic, psychological, or sociological investigation may be compromised, or enabled, in ways determined by those associations the concept "rational actor" carries. Another example is "entropy," which is a concept that has come to take on a multitude of meanings from heat dissipation to time itself. The fundamental claim is that we are concept users and that concepts play a substantial role in navigating our lives.

Conceptual engineering is an emerging interest in academic philosophy that asserts that we need to begin engineering concepts (Burgess, Cappelen, and Plunkett 2020; Eklund 2017; Scharp 2013; Burgess and Plunkett 2013). To engineer is to create, augment, and/or build in accordance with some goal or desire. The engineering of our concepts then means the intentional improvement and creation of those mechanisms which enable our abilities such that we can be more effective actors. We can thus understand conceptual engineering as a corrective to the haphazard way we go about acquiring our concepts—a mostly unintentional activity. We pick up concepts from peers, media, and culture generally. They are often confused and rarely planned. Education and schooling can afford us greater control when it comes to the refinement of our concepts. However, curiosity still exists as to whether greater mastery is possible, and with it, superior control. It seems strange that such necessary machinery has so little planning involved and even less method. The question boils down to: can there literally be an engineering of concepts?

To engineer at all, the engineer must understand what is to be engineered. While conceptual engineering is foremost a normative project, claiming that some concept is superior to another for some end, it also demands a serious investigation into the descriptive side of its materials. Therefore, conceptual engineers must additionally ask: what are concepts? The theoretical weight 'concepts’ hold, given the sheer abundance of functions we have attributed to them throughout the centuries, is immense.[1] A satisfying theory as to the nature of concepts is thus in demand, and has been attempted by recent philosophers along with psychologists and cognitive scientists (Fodor 1998; Brandom 2009; Barsalou 2012; Blouw et al. 2016; Thagard 2014; Millikan 2017; Murphy 2004; Prinz 2002; Churchland 2013; Carey 2011). While this means there is much to work with, the abundance has unfortunately seldom yielded convergence (Churchland 2013; Brandom 2009; Machery 2008). We are left with an immensely tempting project (conceptual engineering)—which promises rewards as grandiose and emancipatory as the labor to bring them about is promethean and herculean. The discontinuity between the theories of concepts may be evidence itself that concepts are beyond our conception in any meaningful way, potentially blocking the promise of conceptual engineering. What are these things we call concepts, and how should we go about engineering them?

While conceptual engineers are by no means certain what concepts are, a received view has emerged. This view takes a language-based approach (Cappelen 2018; Isaac 2020).[2] As it stands: words are the units—the materials of conceptual engineering—and some are happy to conflate words with concepts.[3] As such, meta-semantic theories—questions regarding the supervenience of meaning—are primary, and of course, range from internalist to externalist commitments. The externalists turn their attention to an outside world of dynamics that determine

[1] Kant is the exemplar here (See: Kant 2018; 'Stanford Encyclopedia: Togetherness Principle'). We can also read Nietzsche's transvaluation of all values as holding normative concepts to be orientational, with an insistence on creating new ones (Nietzsche 2003). A third example is Rudolph Carnap, whose theory of 'explication' suggests a refinement of concepts for greater scientific purposes (Carnap 1950). On the continental side, Gillies Deleuze, along with Felix Guattari, built a meta-philosophy around the idea of the philosopher as a concept creator (Deleuze and Guattari 1994). Toward the end of his career, Thomas Kuhn's work explicitly held a 'post-Darwinian Kantianism' that suggested concepts were 'a prerequisite to having beliefs' (Kuhn 1990). The dual insistence on concepts as orientational mechanisms, and the importance of their cultivation, is not new—and it transcends philosophical divides; for a collection with more examples, See: Margolis and Laurence 1999.

[2] I believe this to be the consequence of 1) Analytic philosophers classically being interested in language, and 2) An influence from cognitive psychology in terms of exemplar and prototype theories of concepts (Murphy 2004; Machery 2008).

[3] Others aren't, such as Herman Cappelen who will soon come to be an immensely important person in this paper.

meaning, while the internalists suggest that meanings supervene on mental states. The shared goal is to explain how words arrive at their intensions and extensions. Once we figure that out, then we may proceed with engineering. Conceptual engineering becomes an activity of fitting the right extensions with the right intensions relative to some problems. Psychologists, too, have their own language orientation and have influenced conceptual engineers (Isaac 2020; Murphy 2004). This angle is one that models concepts as exemplars and/or prototypes. An exemplar is a typical instantiation of some object in the world enculturated and ingrained in cognition. We may think the exemplar of a dog to be that which all other dogs are compared. A prototype is fuzzier and organizes concepts based on resemblance. In this case, a hammer might end up classified as a more typical tool than a wrench. A question emerges here: what are exemplars and prototypes? How are they grounded in neural activity and brain anatomy? It seems odd ontologically.

There is a symmetry between philosophers and psychologists who adopt these language orientations: they both approach concepts as categorizations. Any engineering of concepts is thus an organizational game—one that strives to address problems by way of recategorizations. To understand something as such-and-such (a categorization) is to know how to situate it within a framework of other information, which in turn affords abilities to think about, apprehend, and navigate with it. To conceptualize is to render something intelligible, and the way you do that shapes the relations you have with the objects, and people, in your life. While this narrows the focus, there is no shortage of controversy—and of course, pessimistic voices. In the words of Herman Cappelen:

> In most cases the detailed mechanisms that underpin particular instances of conceptual engineering are too complex, messy, non-systematic, amorphous, and unstable for us to fully grasp or understand (Cappelen 2018).

This is known as Cappelen's "inscrutability" challenge to conceptual engineers, and he uses it to bring into question both externalist, along with internalist, views of concepts. In this paper, I will be interested in how both challenge internalist views of concepts, but the overarching program is not one that discriminates between internalism or externalism. As I will show, concepts are things that are very much internal, but they, too, are acted upon by external forces. I provide a framework that is grounded in neurological activity and brain anatomy. My project is one that is substantially different from Cappelen's and language-oriented conceptual engineers. I do, however, provide a place for them.

While one may grow disheartened by "inscrutability," Cappelen follows it with another hit: "lack of control." The additional challenge is precisely what it sounds like: it follows that the abysmal complexity, and messiness of the situation, obscures any attempts to control our conceptual repertoires. Once again, in his own words:

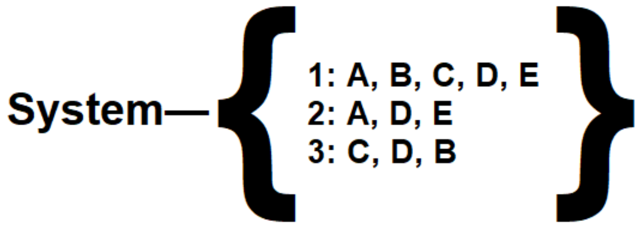
> So there is no shortcut from Internalism to Control. Internalism makes it just as hard to get control. Or, put in terms of a challenge: an argument is needed a) that there are inner states that are scrutable and under our control, and b) that meaning or concepts supervene on those inner states. However, even that isn't enough. Even if you think (a) and (b) are true, you do not have a position that guarantees scrutability or control. You still need (c) the determination relation from the supervenience base to be scrutable and within our control. As far as I know, no internalist has ever even tried to argue for the conjunction of a-c' (Cappelen, 2018).

Cappelen's challenge gets at something I find extremely deep and true about the nature of any conceptual engineering project. I agree with Cappelen that whatever mechanisms we tentatively refer to as 'concepts,' and those that implicate them, are largely beyond us—and that for the most part, we are to their whim. I do not believe we are the central planners or intelligent designers of concepts that we would like to be—instead, a situation persists that has us condemned to the slow churn of culture, along with abysmal neural network operations, to create, and arrange, our conceptual resources.

I accept this as my starting point, and as such, begin immediately moving toward more optimistic territory that explains how complexity science methods can yield scrutability—and someday control. If conceptual engineers are condemned to messy, amorphous, and complex materials, then we should look toward those who have had success with messy, amorphous, and complex materials (Bechtel 2010; Fieguth 2017; Page 2011; Wimsatt 2007; Mobus and Kalton 2015; Clemens 2013).[4] What I mean is that the conceptual engineer should consider the methods of complex system science. Here, 'complexity' can take on useful and even desirable associations. Complex systems possess various functions due to their 'messy' natures: 'redundancy' being one such example. A complex system is said to feature redundancy when multiple parts serve the same function. In other words: if one component fails, then another can pick up the slack, as a single function can be activated by various means.

To a degree, the stability of a complex system is determined by the diversity of its components (Page 2011). Take "A" to be our necessary function for the survival of our complex system, and take "1," "2," and "3" to be our denizens populating our complex system. All denizens have various functions (A, B, C, D, E), with denizens "1" and "2" sharing the function of "A" (this would be the redundancy). For instance[5]:

[4] The Santa Fe Institute is probably the best example. They provide multimedia introductions along with cutting edge research on complex systems.
[5] This model being a variation of the model provided by Page 2011.



If some external influence begins to act upon our system and manages to kill denizen "2" (perhaps an invasive species), we know that the system will survive as denizen "1" also works to produce the function "A." When we inquire further, we learn that the reason why the external influence failed to kill our system was due to the function "C" being present in the denizens "1" and "3," in which case "2" was left vulnerable due to it lacking the function "C." We can conclude that this system is strengthened by the diversity and heterogeneity of its denizens.

If enough redundant processes are present, we can then say that the system is 'resilient' and 'robust,' as it can endure a range of possible influences both internal and external to the system (Page 2011; Jan 2005). In other words: redundancy supplies a series of fail-safes to systems, in turn increasing their fitness and survivability due to the multiple ways necessary functions can be activated (Wimsatt 2007).

Our theories and models, much like the objects they approximate, also exhibit these properties. This amounts to an essential symmetry between map and territory, drawn out by the philosopher William Wimsatt.[6] Here, robustness acts as a confirmation-based epistemology, where the more avenues a scientist has for detecting a phenomenon, the more likely the phenomenon is

[6] Wimsatt frequently comments on the famous physicist Richard Feynman's notion of 'Babylonian Science,' where it is claimed that "the fundamental principles of nature are remarkable because they are derivable in multiple ways using multiple different assumptions" (See: Wimsatt's essay in Lissack 2014), Feynman's example being the varied ways to arrive at the inverse square law (See: Feynman 2013). Even physics can be messy.

to be believed.[7] In other words, a scientific theory (or model) can be more or less robust as it manages to survive a range of possible falsifications/verifications. In Wimsatt's own words:

> We feel more confident of objects, properties, relationships, and so forth that we can detect, derive, measure, or observe in a variety of *independent* ways because the chance that we could be simultaneously wrong in each of these ways declines with the number of independent checks we have . . . Indeed, if the checks or means of detection are probabilistically independent, the probability that they could *all* be wrong is the product of their individual probabilities of failure, and *this* probability declines very rapidly (i.e. the reliability of correct detection increases rapidly) as the number of means of access increases, *even if the means are individually not very reliable* (Wimsatt 2007, 196–197).

Diversity, heterogeneity, redundancy, and robustness are not precisely the dismal descriptors used by Cappelen to explain the conceptual engineer's predicament; they nonetheless paint the same picture: 'messiness' preserved. The point being, they do so while replacing a sense of doom with a sense of opportunity. If our concepts are like our theories and models of science, along with the objects they pursue ('messy,' complex, heterogenetic), and there is an opportunity in this, then perhaps the conceptual engineers can also exploit the 'messiness' of concepts. By inverting Cappelen's challenge, I present complexity as a feature that, if used, can deliver methodological maneuvers for the conceptual engineer, such as modeling what concepts are and what they afford in different contexts. I provide a model that does this while demonstrating how we may someday engineer concepts with greater responsibility. In other words: I am interested in blazing a path from an internalist view of concepts to scrutability and control by way of embracing complexity.

[7] For another voice on this subject, See: Weisberg 2013, 156–170.

How engineering and control are possible must now wait, as some machinery must first be constructed.

Part 1 presents concepts as systems (concept-systems). This section is ontological and is the foundation upon which the rest of the paper rests; Part 2 provides a model of concept-systems, demonstrating the importance of scientific modeling for 'scrutability' in light of complexity; Part 3 concerns a hypothetical conceptual engineering project using the model, demonstrating how one may someday go about employing the model, and making complexity work for conceptual engineers; Part 4 is the conclusion, which returns to 'Inscrutability and Lack of Control.' It considers some additional problems and suggestions for further steps to be taken along the path I am spearheading. This work is linear, and I believe it needs to be read in order.

Part 1: Ontology

This section is tasked with the question: what is to be engineered? If the complexity of concepts is to be exploited, then we must know the materials we are working with. We need an ontology of concepts that is system-oriented. While I believe that language is the prime mover in terms of conceptual engineering—we manage to engineer our concepts with intention and accuracy through re-arranging categorizations—I do not believe concepts should be conflated with words, and that these strict language-ontologies fail to access deeper complexities inherent to concepts where other operations beyond (but also including) linguistic categorization are taking place. Beneath intensions, and extensions, along with the exemplars and prototypes that bubble to the surface, there exists an abyss of complexity untapped by any conceptual engineers I am familiar with. It comes down to the fact that any concept we possess is also a collection of affects, memories, simulations, associations, and other information concerning some target object. These reservoirs of information are contained in brain networks trained on different inputs (such as the sensory-modalities). Concepts pull from depths well beyond linguistic categorization to accomplish whatever we as cognitive agents need to accomplish. It is this complexity that I preserve and find utility in. Thus, this section will provide a space for language-based theories of concepts, and focus on those extra-linguistic components to concepts that are not encoded symbolically. As such, I aim to provide what I take to be a more informed view of the ontology of concepts and demonstrate how such a view carries with it a greater utility for an engineer, beyond (but also including) the language orientation.

So what then do we find when we peer into the abysmal operations of the brain? I believe we find representational devices, and in this paper, 'representational devices' will refer to those multi-modal (of the sensory brain regions) and amodal distributed neural structures which, when activated in parallel, form what I will call a 'concept-system.' What I mean by this is that representational devices are the parts in which concepts are comprised. The key is that there are different kinds of representational devices. Their discoveries have been the product of research accompanied by relentless debates throughout the history of psychology, neuroscience, and the philosophy of mind (Clark and Millican 1999; Eliasmith and Bechtel 2006; Machery 2009; Murphy 2004; Ramsey, Stich, and Rumelhart 1991). More recently, less combative, more inclusive views have emerged that amount to a kind of pluralism which includes symbolic, connectionist (sub-symbolic), and simulation-based representations (Barsalou 2012; Eliasmith 2013; Thagard 2019).[8] It is a convergence that favors multiplicity over elimination by using the smorgasbord of representations history has uncovered in an attempt to explain cognitive activity with a completeness otherwise lacking. As claimed by Lawrence Barsalou, "Successful theories [of concepts] in the future are likely to integrate all three frameworks [symbolic, connectionist,

---

[8] Edouard Machery's 'Offloading Hypothesis' is also continuous as it provides roles for both amodal and modal representations in conceptual processing (Machery 2016).

simulations] into a single system. It is unlikely that theories implementing only one or even two of these approaches will succeed. What each approach offers appears essential to the human conceptual system" (Barsalou 2012). This amounts to a consensus that holds the brain to be a series of networks updated by different kinds of sensory information. The various networks then go on to conspire together (forming a binding of representations—a concept-system), with more or less capability to surmount a given challenge.

1.1 Symbolic Representational Devices

To begin, let us consider symbolic representation. Let us carve out a place for the language-oriented conceptual engineer. Presumably, this representational device would be the encodings of the categories such as intension/extension and prototype/exemplar. Historically, variations on this theme have gone by many names, including 'the theory-theory,' 'the symbolic,' 'physical symbol systems,' 'semantic memory' and the anagram 'GOFAI' (Good Old-Fashioned Artificial Intelligence). The brain (hence, 'physical') was taken to be a kind of computer (possessor and manipulator of symbols). With cognitive operations then sharing a symmetry with computers, computers became the model for thought, and both were made continuous in a singularly harmonious and satisfying explanatory framework.

At its core, this kind of representational device is one which is language-like, consisting of symbols manipulated by the rules of logic.[9] Statements are the symbolicist's representation—and are held as being stored in amodal brain regions. In the words of Lawrence Barsalou, "The resulting representations have the flavor of detached encyclopedia descriptions in a database of categorical

[9] To situate historically, we can think of symbolicism as the continuation of the Fregean tradition into computers, brains, and cognitive psychology (See: Frege and Beaney 1997).

knowledge about the world" (Barsalou 2012). By "detached," Barsalou means that they are "amodal," which means not dependent upon sensory modalities.

By employing a combinatorial logic, brains could be understood as taking simple representations and creating more complex representations (Fodor 1998; Werning, Hinzen, and Machery 2012). This explains how brains go beyond what they acquire from sensory input, granting abstract reasoning. Take, for instance, the category "PET" and the category "FISH." If we consider the conjunction "PET FISH," we can imagine new inferences being afforded, particularly those concerning ownership, responsibility, and companionship with a fish. It is this kind of categorical knowledge, and the various combinations, which made symbolic representation powerful.

Anatomically speaking, the cerebral cortex appears to be related to language and symbolic processing. By way of FMRI, we can observe, in real-time, structures in the cortex being activated by natural language. These structures are known as the 'semantic tiling of the cortex,' and they are distributed along both hemispheres (Huth et al. 2016).[10]

Symbolic systems are not without their problems—perhaps the most explicit issues arise as results of their categorical and nondynamic nature. Symbolic systems are traditionally self-contained: a programmer creates sentences and rules, and this computer imagery comes with some consequences. What is produced is a function of what is intentionally installed. The world is immense, and to build intelligences that navigate it demands categories challenging to produce artificially. Imagine, for instance, trying to program those categories people use to harvest, refine, and create with lumber. The necessary categories range from the tools to deforest, to those pertaining to environmental laws, perceptual functions to recognize, and motor functions to chop

[10] Barsalou comments on this as well, suggesting that while this is "impressive research," given the nature of FMRI research, we are not provided with a satisfying "algorithmic" explanation (See: Barsalou 2017b).

the wood, among numerous others. This quickly becomes abysmal to the programmer—be it human, culture, or nature[11]. From this vantage point, the abundance of the world demands an abundance of sentences and categories well beyond (if at all possible) the programmer (and possible memory). People also manage problems that do not deal with set definitions, suggesting an intelligence altogether different than a clean symbolic one. While there are reasons and evidence to suggest we are partially symbolic in our representational capacities, both of these problems further distance us from the symbolic conception in any totalizing sense.

1.2 Connectionist Representational Devices

A more growth-oriented theory of representation would emerge to address the problems of symbolicism. Rather than having a computer scientist program, devices were created that could learn (Rumelhartet et al. 1986). This demanded a shift from the symbolic to the more biological image of neural networks. Through the advantage of learning, progress was made in terms of object recognition and categorization (Krizhevsky, Sutskever, and Hinton 2012). Neural networks could be 'trained,' allowing for a more dynamic representational device, allowing for problems difficult to represent symbolically to be navigated. This success was due to the structure of neural networks (their 'architecture' as memory) and probabilistic/statistical operations performed on various said structures.

In a biological system, a neural network is a collection of neurons connected in various ways by dendrites. Each cell fires an electrical signal along a dendrite, which in turn either activates or does not activate another cell downstream of it. There are different ways neural networks

[11] I imagine this would also be a serious issue for a conceptual engineer.

represent—including the rate within which they fire and the patterns they fire in accordance with—properties emergent from spiking and inhibitory/excitatory links (Thagard 2019).

If you scale this process up to over 85 billion neurons, the connectionist believes you can attain a human brain. Like the symbolic representations, connectionism, too, has its own shortcomings. Barsalou mentions that even in light of some connectionist models being used to reproduce limited symbolic operations, "So far, connectionism has not succeeded in convincing the cognitive psychology, cognitive science, and cognitive neuroscience communities that this approach explains the basic conceptual operations of propositions and productivity" (Barsalou 2012). This effectively makes connectionist representation a yin to the symbolist representation's yang, hence the insistence by some to synthesize (Barsalou 2012; Eliasmith 2013; Thagard 2019) rather than eliminate (Churchland 1992; 2013).

1.3 Simulation / Modal Representational Devices

The third example breaks from traditional representations in some interesting ways, and it finds its genesis in the embodied cognition paradigm, which holds its representations as being grounded in sensory-motor modalities heavily determined by the body and its interaction with its environment (Barsalou 1999, 2003, 2008, 2012, 2016). As the story goes, the networks which comprise the sensory-motor modalities (e.g. perception, tactile, auditory, olfactory, motor) store information regarding what has been interacted with through the senses.[12] When an agent comes across future

[12] Barsalou explains that this does not make 'simulation/embodied/situated' views strictly an empiricist matter, as he claims it is entirely possible for there to be nativist elements such as architectures in the modalities which are arranged by genetics (Barsalou 2012).

instances of whatever environmental contexts have previously arranged it, these architectures can be reactivated such that a 'simulation' is possible (Barsalou 2012).[13]

A simulation is a patchwork of different sensorial modalities activating in parallel, all converging on a category in what can be thought of as almost a re-living of the situation (Barsalou 2016).[14] If the category is "STREET CAT," then we may find activations such as the perception of "BRINDLE," the olfactory sensation of "PUNGENT," and the tactile sensation "SOFT." These varied representations afford different possible inferences and predictions about cats, granting possible actions (Barsalou 2016).

Simulations do not happen in vacuums, and are 'situated' within an environmental context, expanding possible inferences to those which concern the category under consideration (e.g. "HOUSE CAT") in a contextual setting (e.g. "LIVINGROOM") (Barsalou 2003, 2016). This is due to environmental context also arranging the architectures of our modal networks (hence the 'situated' character of simulations). Research supporting this has discovered that various modalities are activated even when there is no explicitly necessary need for them, such as those dealing with the motor activation of grasping a hammer when we perceive a hammer. This explains a kind of priming for our motor system and can explain predictive processing (Barsalou 2016).

1.4 To Take Inventory of our Representational Devices

(1) Symbolic Representations

- Language-like, computational, amodal, massively distributed along the cortex. This is likely where intensions, extensions, and exemplars, prototypes lay anatomically.

[13] 'Offline' processing is also possible, meaning simulations can be activated without contact with the initial objects (Barsalou 2012).
[14] Although, this is largely an unconscious activity (Barsalou 2016).

- Afford propositional representations in terms of categories, problem-solving, when problems are explicit.

(2) Connectionist Representations

- Sub-symbolic, probabilistic/statistical, massively distributed.
- Afford object recognition, prediction, learning from the environment (adaptation), handle ambiguous problems well.

(3) Simulation / Modal Representations

- Simulation-based, distributed across sensory-motor modalities.
- Afford simulations that grant inferences to be made regarding various objects in various contextualized settings, allows imaginings, affords situated action, and affective states.

1.5 The Benefits of Multiple Representational Mechanisms and a Problem

It follows that an orchestration of representations would amplify one another, affording a more comprehensive range of cognitive capacities than any of them do in insolation. When considering a bicycle, for example, we need an assortment of inferences, predictions, affects, and categorizations to properly use the bike. This could demand motor and tactile simulations, granting the necessary extension and flexion movement to pump the pedals. The bike ride also requires perceptual networks and simulations that afford the identification of objects along the trail, with the necessary inferences and predictions about them. Our aggregation must also include the rules of our local park to be activated in semantic memory, such that we may conduct ourselves with the proper etiquette when passing riders and pedestrians. In other words: a complete cognitive agent is explicable through the inflation of our representational devices.

Notably, the agency various representational kinds grant extends to more sophisticated processes such as science itself, and the philosopher Paul Thagard develops this point in his ‘combinatorial conjecture.’ In his own words:

> Many scientific discoveries can be understood as instances of conceptual combination, in which new theoretical concepts arise by putting together old ones. Two famous examples are the wave theory of sound, which required development of the novel concept of a sound wave, and Darwin’s theory of evolution, which required development of the novel concept of natural selection. The concepts of sound and wave are part of everyday thinking concerning phenomena such as voice and water waves. The ancient Greek Chrysippus put them together to create the novel representation of a sound wave that could explain many properties of sound such as propagation and echoing. Similarly, Darwin combined familiar ideas about selection done by breeders with the natural process of struggle for survival among animals to generate the mechanism of natural selection that could explain how species evolve (Thagard 2014, 109).[15]

We, too, can imagine now that the components of concepts such as ‘sound,’ ‘wave,’ and ‘selection’ are all comprised of different representational kinds cultivated through a mixture of interactions with and conversations about various objects and environments. Along with the methodological implications, we can begin to see how scientists and philosophers are built from the everyday by way of the combined activation of various networks containing various representations, trained by various sensory inputs. Now that we have a taxonomy of representations, we have the stuff of concept-systems, and as such, we must ask how these different representations work in tandem—how do they form concept-systems? How the brain manages to accomplish this is up for debate,

---

[15] Thagard provides more examples of his ‘combinatorial conjecture’ in the history of science (See: Thagard 2013, and for further elaboration, See: Thagard 2012).

and there exist several possible mechanisms to explain it (Barsalou 2017a; Thagard 2014, 107–140).[16]

1.6 Semantic Pointers as Solution

A tempting mechanism that weaves representations in a way commensurate with what I am attempting here is the semantic pointer, as proposed by Chris Eliasmith and expanded upon by Paul Thagard, along with others (Stewart, Choo, and Eliasmith 2014; Blouw et al. 2016; Eliasmith 2013; Thagard 2019). A semantic pointer is yet another kind of representation that is symbolic. As described:

> In its most basic form, a semantic pointer can be thought of as a compressed representation that captures summary information about a particular domain. Typically, such representations derive from perceptual inputs. An image of an object in one's visual field, for instance, will initially be encoded as a pattern of activity in a very large population of neurons. Through transformations . . . however, further layers of neural populations produce increasingly abstract statistical summaries of the original visual input. Eventually, a highly compressed representation of the input can be produced. Such a characterization is consistent both with the decrease in the number of neurons found in later hierarchical layers of the visual cortex, and with the development of neurally inspired hierarchical statistical models for dimensionality reduction . . . Analogous representations can be generated in other modalities such as audition and sensation (Blouw et al. 2016).

[16] There are a few mechanisms known to complexity science that manage to organize various members of a system such that they labor in tandem, and with aggregative effect, each being more or less applicable to the brain. These include synchronization and swarming (Strogatz 2004; Bonabeau, Dorigo, and Theraulaz 1999; Mobus and Kalton 2015). It may prove fruitful to explore these mechanisms as they relate to representations, brains, and conceptual engineering. In many of these examples, self-generating systems are typical. In the case that there is no weaver—no top down creator of concept-systems as semantic pointers may be describing—then conceptual engineering becomes trickier, but not impossible. We do in fact manage to engineer such complex systems.

Semantic pointers emerge when the brain builds compressed representations of representations. In turn, semantic pointers are relatively simple mechanisms compared to the underlying complexity stored within those subordinate networks that they activate. The full gamut of representational resources across the brain is then reachable through decompression (i.e. 'pointing'—another adjective they use is 'regenerate,' as to say semantic pointers regenerate representations).[17] The degree to which a semantic pointer activates those lower-level representations and their greater complexity is relative to the problem at hand and how much experience the networks have with such a problem and its corresponding objects. It all pivots on how much information is needed to surmount some challenge. Thus, more significant issues demand greater representational depth, as pulled from those subordinate representations housed in the modalities—semantic pointers explain this process.[18]

> They [semantic pointers] can account for symbolic processes, perceptual simulations, and a host of other functions all centered upon a single object class. In other words, they can act as a summary representation of a category of things in the world, which is precisely what a concept is often taken to be" (Blouw et al. 2016).

It is important to stress that C, Eliasmith, P, Thagard, etc. do not conflate semantic pointers with concepts, but rather semantic pointers are the means within which concepts arise. Semantic pointers build concepts out of the materials they have to work with: the various representations encoded in the brain. Concepts are then the bindings of multiple representations across modal and amodal regions relative to objects and problems. Concepts are thus systems, and are in statu nascendi, with semantic pointers potentially serving the function of a midwife. The brain comes

---

[17] This is elaborated further in terms of convolutional neural networks, and recursive binding (Eliasmith, 2013)

[18] For greater detail as to the success of syntax reconstruction via semantic pointers, see: Stewart, Choo, and Eliasmith 2014. For greater detail as to the success in terms of prototype and exemplar models of concepts via semantic pointers, see: Blouw et al. 2016.

to be understood as a thing that builds concept-systems relative to some object(s) or problem(s) it must navigate. The properties of these concept-systems afford the options to be taken, ignored, or approximated. We are as capable as our concept-systems, and our concept-systems are as capable as their representational devices.

While Chris Eliasmith has managed to implement his Semantic Pointer Architecture (SPA) in an extensive model of the human brain, achieving orchestrated multi-representation success, it is additionally important to note that semantic pointers have not been verified in biological brains. They currently exist as a postulated entity: one possible way, among few others, which explain how coordination is possible for modular representations. While the case may be stronger for representation (and they, too, are still controversial) than it is for their organization (orchestration into complex-systems), the binding of these representations into systems in biological brains is adduced (more on this in the conclusion). This is a move I make that might not be satisfying; for now, I continue moving. With further transparency: I am also presupposing that when concepts are fashioned out of the armamentarium of representations that they do so with enough consistency, relative to the challenges of the world, such that concept-systems are also generalizable, meaning that they can be modeled, predicted, and finally engineered.

1.7 An Additional Word on the Activation of Different Representations

It is important to note that these various representations can all play the roles of activating a concept-system. Semantic pointers are built with perception being the initiator of the process. Still, any sensory input presumably can also act as an initiator, triggering a semantic pointer, which then weaves together a concept-system. We understand cognitive agents then as being comprised of many sensors, all of which can get the ball rolling on various concept-systems. This is an important point I will develop in greater detail later, briefly: this is how one may engineer something like a perceptual representation by way of words. Imagine an illusion that is not intelligible until it is situated verbally within some context.[19] Modern art often suffers from this same problem: by lacking an explicit narrative, some people are turned off. The piece becomes troubling; it is that thing which cannot be fitted to a concept-system—an intelligible framework. This can be uncomfortable, and some people may desire a curator or artist to explain the piece—to engineer what it is they see, to help them situate the piece in their understanding, to entangle the work with some concept-system(s).[20]

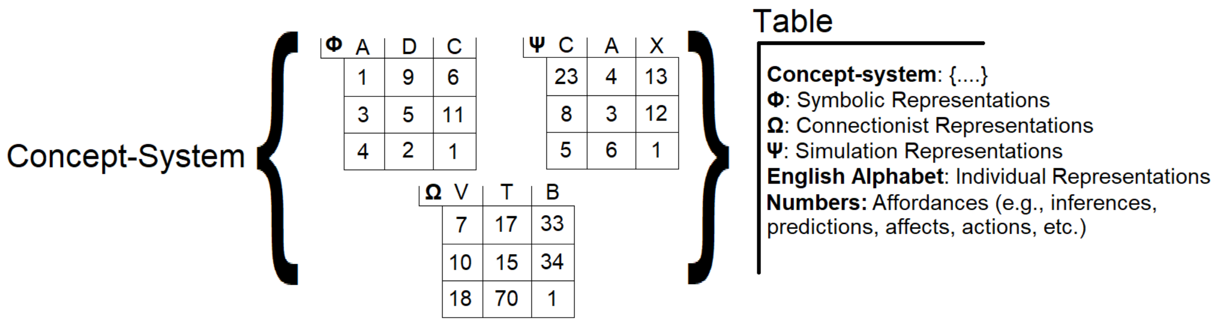
Part 2: Complexity and the Modeling of Concept-Systems

I have explained different representational kinds, and I assert that they are likely being assembled into concept-systems through some mechanism (possibly semantic pointers). I now have to explain how complexity properties emerge, and I will thus be providing a model of concept-systems to do so. The following section takes this as its target.

Like conceptions of concepts, measures of complexity are plentiful (Ladyman et al. 2013). What is less controversial is the ubiquity, and generality, of some properties of complex systems, such as robustness, diversity, and redundancy (Wimsatt 2007; Page 2011). As the introduction claimed, these properties enable the conceptual engineer rather than hinder or block indefinitely.

[19] To see the duckrabbit as a duck or a rabbit demands the concept-systems "duck" and "rabbit."

[20] Much of this is continuous with 'the theory-ladenness of observation' as articulated by 20th century philosophy of science, where it was said that observation is partially determined by theories. This is due to 'theory,' like 'concept,' being orientational (preceding and guiding observation). We can replace 'theory' with 'concept' and capture much of this story, with potentially new and interesting consequences. Paradigm change is also made explicable here, as scientific revolution can be said to be a product of a change of orientational mechanisms (concept-systems). I suspect a 'concept-ladenness of observation' to be a fruitful pursuit. For more on the theory-ladenness of observation, See: Hanson 1958; Kuhn 1962; Popper 1963; Feyerabend 1975; Hoyningen-Huene 1993; Hintikka 2007.



Let us return to the model from the introduction, but add another dimension (affordances), and fit our conception of concepts to it. What is shown here is a hierarchy demonstrating a concept-system. This model is also a map of where a conceptual engineer may intervene in a concept-system.[21] At the top of the hierarchy, we have a concept-system, which decomposes into the three kinds of representations considered: symbolic, connectionist, and simulation. Each representational device then goes on to further decompose into individual representations belonging to each representational kind. This time around, the selective pressures acting on the concept-system would be the numerous problems the world demands cognition to surmount. To meet a challenge, the concept-system affords several inferences, predictions, affects, actions, categorizations, etc. Either the concept-system rises to the occasion, or it does not, as a function of what the representation(s) afford(s).

2.1 Robustness and Redundancy

---

[21] Another way to think of this is via Herbert Simon's 'near decomposability,' where it is said that agents embedded in complex systems break the world down into workable chunks, such that mechanisms can be understood reductively and hierarchically, allowing complex systems to be navigated accordingly. The sort of complexity framework for conceptual engineering that I'm inching toward is a kind of extension of this into more contemporary theories of mind, where 'concepts-systems' function as the decompositions of the world, which in turn decompose further into the various representational kinds (For Simon on complexity and hierarchy, See: Simon 1962. For more on near decomposability, See: Simon 1996; Wimsatt 2007. For Simon on complexity, mind, and symbolicism, See: Simon 1996).

Using the above model, let us set the concept-system to "DOG" and imbue our cognitive agent with modal representations: where "ΨX" is the simulation of a dog wagging its tail in a pleasant setting, and "ΦA" is the symbolic information that 'dogs are happy after they eat.' Now, let us embed our agent in a world where a dog is present. The concept-system "DOG" becomes activated by way of ventral and dorsal activation (perception). Let us set this to "ΩV18." Our agent notices the dog has eaten and is wagging its tail; hence the agent is delighted rather than frightened, due to the joint representations "ΨX" and "ΦA" both affording "1," which is the affective state, delight (this is the redundancy). Thus, because of "ΨX1" and "ΦA1" both being present, we can say our agent's concept-system "DOG" is robust relative to the challenge of reacting appropriately to the dog. If one representation, say "ΨX," were inaccessible, we could still count on "ΦA" to afford a desirable reaction. The inaccessibility could be due to external circumstantial factors such as an inability to see the dog wagging its tail. The more affordances converge between representations, the greater chance we have to act with success. We can then define concept-robustness as the resilience of a concept-system relative to a problem as a function of redundant representations and their aggregation.

2.2 Robustness and Redundancy in Multi-Model Usage

In less ideal, more real-to-life circumstances, multiple concept-systems are activated and working in tandem, meaning that their representations constrain and enable one another beyond what is afforded by singular concept-systems, such as what our "DOG" example demonstrates. This effectively means concept-systems interface with one another, substantially increasing the complexity (and affordances) present in a cognitive architecture due to interactions between representations of different concept-systems. In other words: if we scale up by including more concept-systems, we again find robustness. This model also applies—we just multiply the number of models in proportion to the number of concept-systems activated relative to some problem.

Now, let us plug in Paul Thagard's example from the previous section to show this in a hypothetical scientific context. You will recall Thagard claiming that "Darwin combined familiar ideas about selection done by breeders with the natural process of struggle for survival among animals to generate the mechanism of natural selection that could explain how species evolve" (Thagard 2014, 109). We can imagine a "Possible Darwin" having had in his possession two concept-systems:

(1) Wild Animal

(2) Domesticated Animal

Contained in 1, our "Possible Darwin" may have modal representations, observations, and auditory experiences of animals engaged in mortal combat—the stuff of useful simulations, on the path toward understanding selective pressures. The concept-system 2 could possess facts born of conversations breeders may have shared with "Possible Darwin," concerning the then inexplicable phenomenon of traits being passed to offspring. If the "wild animal" and "domesticated animal" concept-systems both afford the inference "heritability" by containing representations of similarity between parent and offspring (e.g. color, size, temperament, etc.), then "Possible Darwin" could see the break in the symmetry between forest and farm—where the former does not have a breeder, and the latter does. Redundancy is present, as the inference "heritability" was accessible by both Darwin's concept-systems "wild animal" and "domesticated animal." This mutually shared function between both concept-systems, combined with the absence of a "breeder" in "wild animals," suggests a mechanism must exist which explains the heritability found in "wild animals." "Possible Darwin" concludes that survival in the wild takes on the functional role of the breeder in civilization; hence the concept-system "natural selection" is created as a mixture of various representations.

2.3 Further Complexities

An engineer's relationship with their materials expands beyond that in which is most apparent. There are rivets, bolts, and I-beams, but the structural integrity of those materials matter too, and determine an engineering project in crucial ways. To only concern oneself with the intensions and extensions (or exemplars and prototypes) of concepts, is to engineer without attention paid to the structural integrity of concepts. We need to measure and to model the inner workings of concept-systems to maximize engineering potentials. Yes, concept-systems do present themselves as categorizations—be they intensions and extensions as the philosopher is interested in, or the exemplars and prototypes the psychologist is interested in. Still, those functions should not be conflated with what the concept-systems themselves are. Categorization is a product of concept-systems but is only one function, among others (inference, prediction, simulation, affect, association, etc.). I suspect deep within the networks I am using to ground 'concepts' we have complex properties and functions beyond robustness and redundancy.[22]

Part 3: Speculations Toward Control

Now that we have a sense of what we are working with: concept-systems, which are bindings of various kinds of representations that enter into coordinated hierarchical and complex relations with one another, we can explore the possibility of control over our concept-systems. Keeping with the

[22] I suspect it is important to look for scaling laws, power laws, swarm mechanisms, feedback mechanisms, emergence, and other non-linear products.

demonstrative style, this section then addresses one question: how can this complex system ontology of concepts, the model provided, and the application of robustness analysis aid the conceptual engineer? I feel it is necessary to be explicit that this section is more speculative, exploratory, and less grounded in the related sciences than the previous approach of mass-synthesis. The preceding parts constructed the machinery—what follows is an attempt to use it.

3.1 How Might Language Influence the Complex Depths of Concept-Systems?

While I do not think concepts are words, I believe that words are still the primary medium of conceptual engineering. Recall back in Part 1.7: I mention that different modal networks manage to activate concept-systems; thus, concept-systems can find their genesis in any sensory inputs being stimulated. This seems to open space for non-language based conceptual engineering, and it is not surprising as we can inform one another through bodily expression. I find this immensely interesting and worth developing, but I am focused on how verbal language manages to engineer. I may show you how to swing an ax, but that might not convey the correct details necessary to advance your felling abilities. It could be that ideally both demonstration and verbal description are used to supplement one another (to activate all those networks which converge on the felling ability). Still, there is a lucidity with language not necessarily found in bodily displays of information. Through words, I can isolate each hand placement on the ax and each angle in every instance used to deliver the blow. Language manages to domesticate the flux of sensory experience, converting it into intelligibility through accuracy. If it can be talked about, we can determine particular options.[23]

[23] In a sense every conversation is a conceptual engineering endeavor, as coordination between two or more speakers is coordination between those concept-systems present in all participants.

How, then, does language engineer those extra-linguistic dimensions to concepts? How does the verbal influence the non-verbal? How can an engineer target those subordinate representations? I do not believe this problem is solved, but more light may be cast on it: consider the various cognitive therapies which use language to aid in changing behavioral habits. Affective states are understood as both other than and an ingredient of our epistemic faculties (Gonzales 2004; Barrett 2017). Our emotions may not be entirely language-based; they might not be at all language-based. Still, we can use language to understand them—to become mindful of them, to categorize them, and to begin, then, altering them by way of changes made to behavior. In cognitive behavioral therapy, we find worksheets that call for patients to articulate their distress (Leahy 2017). Once this is accomplished, the less than favorable behaviors are situated in an intelligible framework; in turn, the patient may then begin experimenting both by way of addressing their disorder and adjusting their understanding. This could mean any number of deeply personal moves are to be made: perhaps the patient learns to recognize triggers and chooses to experiment with exposure; maybe it is best to reason away some paranoia. In such cases, language and categorization are weaponized against mental illness, and more descriptively, against those underlying learned behaviors that are the affordances of non-linguistic representations. Rational emotive behavior therapy puts even more explicit emphasis on recategorization as a therapeutic intervention, as it calls for "those who are depressing themselves about unfortunate events" to "repeat" replacement claims such as "it's bad but it could be worse" to "themselves often and forcibly" (A and D Ellis 2002). One can imagine most of these activities have corresponding reductions to intensions/extensions and prototypes/exemplars.[24] As previously mentioned in section 1.3: in a study conducted by Diane Pecher, René Zeelenberg, and Lawrence Barsalou,

[24] In the case of rational emotive behavior therapy the concept that is "bad," causing dismay, can have its intension changed to "it's something bad, but it could be worse" in a way to minimize the cognitive distress.

different sensory modalities exhibited a measurable "switching cost" when subjects were presented with words that activated different brain modalities. One such example used first was "BLENDER-loud," which activated the auditory modality, then "CRANBERRIES-tart," which activated the taste modality, and finally "LEAVES-rustling," which reactivated the auditory modality with a delay (Pecher, Zeelenberg, and Barsalou 2003). While this study was after the existence of switching costs, which were found to occur when one activates different modalities in succession, this is also evidence that words themselves can activate sensory modalities (and their component non-linguistic representations). This is an immensely important point for the conceptual engineering of concept-systems, as it means that our words and categorizations have an efficacious relationship with those non-linguistic subordinate representations—language becomes a controller or means within which the verbal can engineer the non-verbal.[25]

In complexity and system science more generally, there is a useful concept-system that captures this kind of relationship between the component parts in a system: we can say that language and those subordinate representations are 'tightly coupled.' It means that all of the parts of a system have strict jobs to be performed, and the operations of one shape the other. This sort of dependency relation means that if you wiggle one, the others wiggle too. In other words: systems that are tightly coupled are susceptible to perturbations throughout, and we may use any particular component as a means to influence the others.[26] By using words that the audition and taste modalities were sensitive to, the switching cost study was able to demonstrate a coupling between

[25] Language-as-controller for the conceptual engineer is something I plan on exploring in a future paper. There very well may be an even deeper relationship between complex system science and conceptual engineering by way of control theory. For more on control theory, See: Abdelzaher et al. 2008. In this case a union of control theory, complexity science (in particular feedback loops), linguistics, and rhetoric may prove to be a path forward.

[26] Of course, then, this is also a feedback mechanism as language and subordinate representations each determine one another. Cybernetics, thus, also has a place in conceptual engineering. For more on Cybernetics, See: Mobus and Kalton 2015.

language and sensory modalities. In other words: language was able to wiggle the non-linguistic representations.

We may use our words to build better habits, which, in turn, create better subordinate representations—better concept-systems. This is indirect, and there may someday exist a more direct conceptual engineering project (perhaps eventually even by way of the augmentation of brain anatomy—brain surgery as conceptual engineering). The point is, we use language as such to engineer deeply ingrained representations in therapeutic ventures. I do not see why it cannot also be used for epistemic ventures, which in the case of conceptual engineering, would amount to using updated categorizations to engineer those subordinate representations which concern knowability (e.g. the categorization of x as such-and-such augments subordinate representations y and z).

It is not much that I am asking for; the conceptual engineer with their language orientation ought to continue conducting themselves as they do, but with the carefulness not to conflate words with concepts, and to begin working alongside computational neuroscientists, complex system scientists, and cognitive psychologists to model concept-systems such that their work on intensions/extensions and prototypes/exemplars can be informed by the complexities underlying (but also including) linguistic categorization. What then would such an engineering project look like? The next section considers a cross-pollination of ideas from conceptual engineering and this complexity orientation, focusing on application.

3.2 A Hypothetical Engineering of the Concept-System "Pseudoscience"

The chief question of the philosophy of science is: what is science? This question saw numerous attempts at answering it in the 20^{th} century, with varying degrees of satisfaction (Popper 1963;

Kuhn 1962).[27] While the question of demarcation is immensely important, there exists another problem that I do not believe has received enough attention: it seems that there should be a concept-system denoting those experimental sciences that are not yet sciences, and that this concept-system additionally should not carry the pejorative qualities "pseudoscience" does. We need new sciences to grow, but are dismissive of sciences that are not sophisticated to our liking. As with the nature of all things: sophistication, detail, and maturity, all demand time to cultivate. Thus, there exists a tension, and we could use a non-derogatory concept-system that denotes a science that is both immature and a desirable work in progress.

Let us consider two ways a conceptual engineer may intervene. As David Chalmers suggests, there is the method of "re-engineering," which is precisely what it sounds like: you fix a concept that otherwise fails to meet some requirement (Chalmers 2018).[28] We may opt to re-engineer "pseudoscience" such that it is more descriptive, particularly in the sense that it manages to dismiss those undesirable sciences while leaving space for desirable budding sciences to continue growing.

Let us imagine a hypothetical engineering project where we have modeled reactions an intellectual community has concerning those things they call "pseudoscience." This may range from the usual categorizations and novelties we have all heard a million times ad nauseam—the mantras often attributed to those who fit another pejorative: "scientism." These responses may differ in terms of anatomical activation and expression, but they all share the commonality of holding pseudoscience as harmful, naïve, dangerous, scornful, and loathsome. We have fitted those affordances to different representations, and we understand the modalities that are becoming

[27] For a collection of recent papers concerning the demarcation problem, See: Massimo and Boudry 2013.
[28] Re-engineering may be the category that contains other conceptual engineering projects such as Sally Haslanger's Amelioration project and Rudolph Carnap's Explication project (Haslanger 2019, 2000; Carnap 1950).

active. We have a hierarchical model of the relevant concept-systems, where we find that, in this case, those representations that produce negative affect states seem to be chiefly responsible for blocking any seriousness being granted to the novice sciences. After years of bashing "pseudoscience," our philosophers of science have entangled their beliefs with their emotions. They are not merely dismissive in their reasons, but additionally in their feelings as their sensory-motor modalities have been programmed to respond negatively to anything that resembles pseudoscience. This means that their concept-system "pseudoscience" pulls from representational reserves of information that were designed evolutionarily to confront threats, in which case our hypothetical philosophers of science are compromised by a mix of cortisol production and amygdala activation. They are unable to suspend their condemnations enough, such that they can allow room for the immature sciences to mature. In which case, the path between here and where novice sciences may blossom is through bypassing (re-engineering) those representations producing problematic affects. This, of course, is challenging as it concerns the problem addressed in the previous section: using language-as-intervention into those non-linguistic depths; but let us supplement this system approach with a re-engineering, as our psychologist colleagues would employ. Let us then consider the intensions:

> *Re-engineering the concept-system*: "pseudoscience"
>
> *The intensions are*: fake science, junk science, inadequate science

We can intervene on this level. Perhaps "inadequate science" is an intension (a categorization representation belonging to the symbolic domain) we decide to engineer out of the concept-system "pseudoscience," as an inadequate science is still potentially redeemable, it just has to be made adequate. There is an implied hope here that is not contained in "fake science" (which misleads and obscures) or "junk science" (which is garbage, useless). We may begin by stating the importance of both encouragements being shown to inadequate sciences and the provision of spaces where they can become adequate. This immediately is shot down by our hypothetical philosophers of science who report reasons that are suggestive of a general hesitance to encourage something that may very well lead to false conclusions and a waste of resources and time. They offer reasons, all of which suggest an uneasiness, and are frequent enough to additionally suggest generalizability. These statements are correlated with the activation of those subordinate representations responsible for anxiety.[29]

Anxiety, in this case, is preventing any amenability; our re-engineering appears blocked. However, since we know these subordinate representations of a particular kind are guilty, perhaps we can address all of them in a single re-engineering. We decide to accommodate those problematic subordinate representations with an insistence on the necessity of experimentation and courage in light of those inevitable anxieties associated with experimentation. This is a classic narrative, and one that has a long history of success in terms of circumnavigating fear. By way of this rhetorical strategy, we come to find that our philosophers of science cease associating inadequacy with "pseudoscience" in the desired ways. The intension "inadequate science" is thus eliminated from the concept-system "pseudoscience." Our novice science manages to evade the categorization "pseudoscience." Our re-engineering was a success.

In this case, we found that the hierarchical modeling of concepts as complex concept-systems, and those subordinate representations that cause anxiety, to be empirically apprehended

---

[29] There may be other biomarkers, too, that determine subordinate representation activation by way of measuring arousal. The idea is to find as many mutually independent checks, as Wimsatt suggests in the introduction, to apprehend as many composite information structures (representations) present in our concept-system. This is how we build our models of concept-systems.

and then re-engineered with responsibility and accuracy. To model a concept-system is to model the relevant information encoded in various networks (representations) and what they do in different contexts. In this case, those modalities associated with aggression responses enabled a liberal usage of the term "pseudoscience" in ways antithetical to desirable concept-system usage.

In the same hypothetical, another group of conceptual engineers modeled the extensions afforded by the concept-system "pseudoscience," but when they did, they found that the extensions were: astrology, numerology, homeopathy, flat Earth, and creationism—none of which are the fledgling sciences we have in mind—therefore, they concluded that it would make no sense to engineer the extensions. The problem lay with the intension. This conclusion was also arrived at through scientific measurement.

Perhaps our hypothetical conceptual engineers decide, in a further attempt to legitimize novice sciences, that a new concept-system is necessary to denote them. A second approach Chalmers provides is "de novo engineering," which in this case would be the construction of a new concept-system with new intensions and extensions (Chalmers 2018). There is, however, an illusory freedom here, because it appears that any new concept has not had anything put into it or growth emerging from it, such that it is easier for the categorical representations (intensions/extensions) in our complex concept-system to mean, and refer, what we want them to. There might be some truth to this, but ultimately the words we choose carry affordances of their own.

De novo engineering: a new concept-system for those immature yet hopeful sciences

Let the corresponding name of this concept-system be: "apprentice science"

The inclusion of “apprentice” may activate the right associations, as apprentices are not people to disdain, but rather encourage, help, and enable. There are already affordances here. These sentiments are, of course, paired with the beliefs that the apprentices are naïve and immature, but the point is that we do not hate our apprentices with the same fire that some may hate sciences that also lack sophistication. In turn, this should then provide a good basis in which the new concept-system, “apprentice science,” can itself thrive, and for it to do the desired work. The intensions for the concept-system “apprentice science” could be: hopeful science, science rich in potential, a science with options.[30] All of these intensions ultimately pivot upon the depth of information to be found in those various representational kinds, and while these particular associations we have with “apprentice” are apparent, associations with other categorizations and concept-systems might not be.

Additionally, there may be information we have that lay dormant about “apprentices,” too. It seems crucial that we begin asking what it is our concept-systems harbor that may not be transparent to us; thus, a database of various concept-systems and how they relate to various contexts seems paramount for de novo engineering.[31] This kind of modeling can inform the creation of concept-systems beyond assertion, and speculation, as to what work the conjunction “apprentice” and “science” does and in what context. Perhaps the way to determine affordances that are dormant given some topic, debate, or challenge, is by way of investigating concept-system activation in different settings and then looking to see what does and does not activate across the

[30] One extension may be conceptual engineering.
[31] The case can be made that a de novo engineering project that invents entirely new words avoids some of this problem.

varied contexts—we then inquire as to why, and isolate whatever mechanism(s) block(s) desired activations.[32]

To illuminate, let us say some of our hypothetical philosophers of science fail to adopt the new concept-system “apprentice science.” They have the representations needed; they even have long histories of apprenticeship themselves, in terms of having once been apprentices and mentors to apprentices. In this case, they use the desired intensions for “apprentice science,” but they fail to include the desired extensions as initially formulated. In other words: they are not applying the status of "apprentice" to the intended fledgling science. This may be as simple as updating the extensions, then, such that they include whatever “apprentice science” is demonstrably worthy of apprenticeship. The difficulty here is not so much our hypothetical philosophers of science, as they have the materials to conceptualize apprenticeship—this machinery is just lying dormant. The difficulty, instead, is in the presenting of an immature science itself, which fails to activate those dormant affordances. We can arrive at this conclusion empirically, and that is an imperative ability for conceptual engineering.

Dormant representations and their desired affordances can implicate robustness too. What, then, does conceptual engineering, as informed by robustness analysis, look like? We can say that if 1) we have modeled robust affordances across various representations contained in a target concept-system, and 2) that they are not being activated, then 3) we may opt to use our language in an attempt to activate those desirable and dormant representations. This, then, would suggest that there are many paths to our desired affordance, which also suggests some are more appropriate (and workable) than others, relative to those individuals, challenges, and the world.

[32] This would be its own robustness analysis, as Wimsatt suggests in the introduction. We would be more confident regarding what information is present, but not active, by determining the activation of that information in a range of contexts. Why is it that someone may love in some contexts and not others? Can we target those instances and engineer what is a dormant love affordance to be more expansive?

Let us return to the model to finish this application-oriented section with a final example:

Concept-System

Φ	A	D	C	B	K
	1	9	6	1	34
	3	5	11	5	7
	4	2	1	8	1

Table

Concept-system: {....}
Φ: Symbolic Representations
English Alphabet: Individual Representations
Numbers: Affordances (e.g., inferences, predictions, affects, actions, etc.)

This time around, we do not have to model the connectionist or simulation depths of our concept-system. We know that it is only one representational kind that needs to be re-engineered (symbolic representations). As such, we can use a minimalistic model—but in this case, there are more symbolic representations that are necessary to incorporate into our model than the usual three we are working with. We see in this model that "1" is an affordance that is redundant across many different symbolic representations, as "A," "C," "B," and "K" all manage to produce that potential affordance. It also happens to be the affordance we are after; it amounts to the action we need some concept-user to employ.

Here is the dilemma: we find in our investigation that while this concept-system is loaded with opportunities to meet the desired affordance ("1"), none of them are being activated. We find that symbolic representation "A" is blocked by some financial incentive; we additionally discover that symbolic representation "C" is blocked by a strange animosity our concept-users share; even more troubling, "B" cannot become activated due to a taboo no one should mention; and finally, "K" fails to deliver the desired affordance "1" because it is logically inconsistent with other claims being made regarding the matter at hand.

What is to be done? Perhaps we engineer such that “1” is activated by whichever symbolic representation minimizes harm. In this abstract demonstration, I am not interested in answering which path we should go with. The real point I am striving for here is that this complex system science approach, when mated with cognitive science, lets us know our options. What this shows, ultimately, is that there is a harmony our concept-systems must enter into with the world in order for us to surmount whatever challenge we need, and that some such harmony is optimal given our desires. We come to understand that we must make a choice, and while this is always difficult, we can at least know scientifically what our choices are. We are only ever some arrangement of our concept-systems and the world, away from what, who, and where we want to be, and there are costs involved in choosing, but scrutability is found here through the modeling and embracing of the complexities present in our concept-systems. Conceptual engineering can become a legitimate engineering science.[33]

Part 4: Conclusion

The conclusion considers some modeling challenges and makes some suggestions as to how they may be addressed. I then finally return to Cappelen’s challenges of “inscrutability” and “lack of control” to see what remains and what I have accounted for.

4.1 Addressing Modeling Difficulties

This paper has thus far presupposed quite a bit. There are no shortages of problems that I introduce by way of addressing those Cappelen introduces. However, I find my challenges less damning than

[33] One can also imagine dynamic epistemic logic, game theoretic, and decision theoretic methods to be of use to the conceptual engineer in terms of which representations to activate/re-engineer/eliminate, etc.

Cappelen's, as the difficulties are transformed from theoretical impasses to challenges of measurement, technology, and interdisciplinary coordination. That seems substantially less difficult, and most importantly: it makes conceptual engineering possible.

One such presupposition is that the normative dimension of concept choice is heavily entangled with the descriptive project of modeling our concepts. This kind of conceptual analysis demands an anthropology/ethnography that records what concepts are up to in various cultures—which is a line of research currently being explored (Stich, Mizumoto, and McCready 2018).[34] My model may function best as a heuristic to demonstrate the necessity of complex system science methodology for conceptual engineering. Still, either way, some model(s) must be employed, and filling in the variables is a massive undertaking. After all, the possible affordances even a single concept-system grants are immensely contextual and multivariate.[35]

Nevertheless, we exist during a data revolution, one with increasingly sophisticated methods for the acquisition and navigation of enormous data sets. One promising avenue for determining the presence of a concept-system is FMRI (Connell and Dermot 2011; Mason and Marcel 2016; Cetron et al. 2019). This research may prove a powerful observation-based supplement to more traditional language-based methods for detecting concept ownership, such as those employed by anthropologists and psychologists: pile sorting, exemplar, and prototype questionnaires (Hung-Wen et al. 2015; Machery 2008; Murphy 2004). One can imagine a trans-disciplinary investigation unified in a search for correlations between categorizations and

---

[34] Steven Stich and Edouard Machery have taken up a cross-cultural project evaluating philosophical concepts. For more, See: https://go-philosophy.com/.

[35] Idealization is a strategy which could be of use. When dealing with large-scale systems, modelers often cherry-pick those variables most useful. This process becomes more sophisticated as the relevant parts and dynamics are better understood, growing relatively austere models at the intersection of usability and abundance. For more, See: Weisberg 2013, 98–113.

biomarkers, such as those neural circuitries I am calling representations. This may very well amount to a naturalization, or grounding, of concepts (concept-systems).

Experimental philosophy (Sytsma and Buckwalte 2016; Lombrozo et al. 2018) may also be a nexus between such research—bridging conceptual engineering with those sciences best suited to determine concept-system ownership, concept-system structure, and contextual affordances. In the simplest form, the question is, and will be: what concept-systems are woven out of various representations with enough consistency and stability, relative to problems, such that our engineerings may be better informed?[36]

4.2 Cappelen

In closing, let us return to Cappelen's challenges and see how this machinery chips away at some of them.

Scrutability: that there are inner states that are scrutable

To which the machinery provided explains: there appear to be different representational kinds the brain harbors in various regions. They are trained by different inputs and store different kinds of information about various objects of the world. These states become more or less active relative to challenges the world harbors. They optimize by way of composing into concept-systems, and those systems determine how capable we are relative to those worldly challenges. FMRI increasingly supports the likely existence of these representations (Pecher, Zeelenberg, and Barsalou 2003; Connell and Dermot 2011; Mason and Marcel 2016; Cetron et al. 2019). One can

[36] Generalizability lays at the heart of my project here, and is a topic that has received much consideration as a reaction to the reproducibility crisis in psychology. For more, See: Barsalou 2019.

imagine these trends will continue as the visual modeling of the brain (and body) continues to develop. This budding verification methodology, supplemented by anthropological and psychological investigations into the operations of concepts across cultures, suggests generalizability, which then means the possibility of modeling and scrutability in terms of scientific measurement. Once we have such models and understand how they relate to various debates, objects, or challenges, we may apply several methods to engineer responsibly (and with control)—complexity science being one such armamentarium. Complexity and messiness are thus not the hallmarks of doom, but rather the grounds upon which we pivot to the tried-and-tested methods of complexity science. Scrutability may demand inter-disciplinary work, but it also appears as a rapidly approaching horizon rather than an impossible destination.

Supervenience: that meaning or concepts supervene on those inner states

Semantic pointers are one possible way concepts supervene on inner states. As I mentioned in a note: there may be other examples worth exploring in complexity science itself. I admit I may not have satisfyingly answered how representations go on to form concept-systems, but I suspect there is some mechanism to be discovered which provides these functions. I appeal to two reasons: a) the likely existence of different networks in the brain which encode different sensory information about objects, forming multi-modal reserves of information to be used (Barsalou 2016), and b) the survival benefit to such coordination of networks. These questions still stand: what is it that reaches down into the depths of the various networks, comprising the brain to retrieve whatever additional bits of information are necessary (if anything does)? How do the multiple networks of the brain conspire? Answer those, and a conceptual engineering that maintains a fidelity to the title 'engineer' is possible. As for semantics, I do not have an answer, but suspect there will be one. I imagine there are deeper possible engineerings to be arrived at if we did know how semantics works, but in the framework I provide here, all we need to know is: that some meanings are associated with some concept-systems, and that they become active given various debates, objects, challenges, etc. In the meantime, massive inventories should suffice.

> *Determination relation*: the determination relation from the supervenience base to be scrutable and within our control

Again, semantic pointers may be that determination relation. There is an increasing body of work laboring in this direction, and while semantic pointers have not been verified in biological brains: they are biologically plausible (Stewart, Choo, and Eliasmith 2014; Blouw et al. 2016; Eliasmith 2013; Thagard 2019). We may not currently be able to control whatever mechanism is responsible for weaving concept-systems, but we might be building them, or something very much like them. Optimism is warranted.

While this paper takes a more internalist route, the inclusion of complex system methods is entirely commensurate with an external world and its influence—we just add variables for those external actors which contribute to the arrangement of our concept-systems. In other words: we embrace concept-systems as open-systems. We are not necessarily hindered by complexity, but can be enabled by complexity: where there is more, there are also more places to intervene—options to be taken. Intervention with responsibility only exists where computability, measurement, modeling, imagination, and boldness are present.

Forests are an obvious way to understand complex systems. To manage them, silviculture and forestry designed a range of hierarchical models at different scales such that the dynamism of, and between, the insects, plants, trees, animals, fungi, and bacteria can be captured (Messier, Klaus, and Coates 2014). From predation models, agent-based models, and simulation, to clear-felling, controlled burns, and tillage, those who manage forests have designed several ways to inject intentional action into complex systems, bridging our normativity and the abundance of the world. None of these bridges were discovered: they were designed, and that is the kind of spirit the conceptual engineers must work in (any engineers for that matter). It is a messy business engineering, and sometimes it demands large scale coordination between various disciplines. There are many materials out there, and some arrangements of them move us closer to a conceptual engineering project in the most literal sense.

Cappelen is correct that we do not have the desired control, but this is because we have not yet overcome our epistemic limitations, and not because of some additional factor. Cognitive science is a wilderness at the moment: its paths are messy products of expeditions and trailblazing—its bestiary continues to grow, but discoveries are emerging that are beginning to allow us to understand concepts as hierarchical, mechanistic, and generalizable, while preserving their contextuality on the hike toward greater control. The question boils down to: what are the mechanisms to be engineered? If answered, new interventions hitherto unimagined are likely to be discovered (such as the examples of hierarchical modeling and robustness I have demonstrated). What are the fault lines, the thresholds, boundary conditions, laws, and feedback loops governing the mechanisms of conceptual structure?

It is premature to suspect control is impossible. In some cases, control is already here and has been. One non-controversial example of intentional macro-level control, albeit a dark one, is propaganda: where the concepts ‘nation,’ ‘people,’ ‘enemy,’ and ‘ally’ are often manipulated with precision, and consistency, by various vested interests. I cannot think of anything more critical given the current political climate of violence and disinformation, than to scientifically model how we are enabled and limited by our concept-systems in particular contexts. How do propagandists manage to turn one group so consistently on another? Is there underlying machinery here that can be used to bring groups together for good, and with authenticity, honesty, and transparency? There are other great epistemic expanses and success thereof; there is a massive demand for conceptual engineering, and our ability to intervene for good requires that we embrace the complexities present in our concept-systems.

References

Abdelzaher T., Diao Y., Hellerstein J.L., Lu C., Zhu X. 2008. *Introduction to Control Theory And Its Application to Computing Systems*. In: Liu Z., Xia C.H. (eds) Performance Modeling and Engineering. Springer, Boston, MA. https://doi.org/10.1007/978-0-387-79361-0_7

Barrett, Lisa. 2017. *How Emotions Are Made: The Secret Life of the Brain*. New York: Houghton Mifflin Harcourt Publishing Company, 72.

Barsalou, Lawrence. 1999. "Perceptual Symbol Systems" in *Behavioral and Brain Sciences*, 22: 577–660.

Barsalou, Lawrence. 2003. "Situated Simulation in the Human Conceptual System" in *Language and Cognitive Processes*. 18: 513–562. 10.1080/01690960344000026.

Barsalou, Lawrence. 2012. "The Human Conceptual System" in *The Cambridge Handbook of Psycholinguistics*, edited by M. Spivey, K. McRae, and M. Joanisse. New York: Cambridge University Press, 239–258.

Barsalou, Lawrence. 2016. "Situated Conceptualization: Theory and Application" in *Foundations of Embodied Cognition: Perceptual and Emotional Embodiment*, edited by, Coello, Yann, Fischer, and Martin. U.K.: Psychology Press.

Barsalou, Lawrence. 2017a. "Cognitively Plausible Theories of Concept Composition." Y. Winter and J. A. Hampton (Eds.) *Compositionality and Concepts in Linguistics and Psychology*, edited by Winter, Y, and J.A. Hampton, 9–30. London: Springer Publishing.

Barsalou, Lawrence. 2017b. "What Does Semantic Tiling of the Cortex Tell us about Semantics?" in *Neuropsychologia*. 105.10.1016/j.neuropsychologia.2017.04.011.

Barsalou, Lawrence. 2019. "Establishing generalizable mechanisms." *Psychological Inquiry*, 30, 220-230.

Bechtel, William. 2007. *Mental Mechanisms*. New York: Routledge.

Bechtel, William, and Robert Richardson. 2010. *Discovering Complexity: Decomposition and Localization as Strategies in Scientific Research*. Princeton University Press.

Blouw, Peter, E. Solodkin, P Thagard, and C. Eliasmith. 2016. Concepts as Semantic Pointers: A Framework and Computational Model in *Cognitive Science*, 40 (5): 1128–1162.

Bonabeau, Eric, Dorigo, Marco, Theraulaz, Guy. 1999. From Natural to Artificial Swarm Intelligence. Oxford University Pretty.

Brandom, Robert. 2009. *Reason in Philosophy: Animating Ideas*. Belknap Press. Harvard University Press. 177–224.

Burgess, Alexis, and David Plunkett. 2013. "Conceptual ethics, I." *Philosophy Compass*, 8 (12): 1091–1101.

Burgess, Alexis, Herman Cappelen, and David Plunkett. 2020. *Conceptual Engineering and Conceptual Ethics*. Oxford University Press. Oxford University Press.

Cappelen, Herman. 2018. *Fixing Language: An Essay on Conceptual Engineering*. Oxford University Press. 72, 158.

Carey, Susan. 2011. *The Origin of Concepts*. Oxford: Oxford University Press.

Carnap, R., 1950. *Logical Foundations of Probability*. University of Chicago Press, Illinois.

Cetron, Joshua, A. Connolly, S. Diamond, V. May, J. Haxby, and D. Kraemer. 2019. "Decoding individual differences in STEM learning from functional MRI data." Dartmouth College, Hanover New Hampshire.

Chalmers, David. 2018. "What is Conceptual Engineering and What Should it Be?". Philpapers.org. https://philpapers.org/rec/CHAWIC-3

Churchland, Paul. 1992. *A Neurocomputational Perspective: The Nature of Mind and The Structure of Science*. MIT Press.

Churchland, Paul. 2013. *Plato's Camera: How The Physical Brain Captures a Landscape of Abstract Universals*. Cambridge, MA: MIT Press.

Clark, Andy, and David Chalmers. 1998. "The Extended Mind" in *Analysis*, vol. 58, no. 1, doi:10.1093/analys/58.1.7.

Clark, Andy, and Peter Millican. 1999. *Connectionism, Concepts, and Folk Psychology*, vol. 2. Oxford University Press.

Clemens, Walter. 2013. *Complexity Science and World Affair*. Albany: State University of New York Press.

Connell, Louise, and Lynott Dermot. 2011. "Modality Switching Costs Emerge in Concept Creation as Well as Retrieval" in *Cognitive Science*, vol. 35, no. 4, doi:10.1111/j.1551-6709.2010.01168.x.

Craver, Carl. 2013. *Explaining the Brain: Mechanisms and The Mosaic Unity of Neuroscience*. Oxford University Press.

Deleuze, Gillies, and Felix Guattari. 1994. *What is Philosophy?*. Translated by Hugh Tomlinson and Graham Burchill, Verso.

Eklund, Matti. 2017. *Choosing Normative Concepts*. Oxford University Press.

Eliasmith, Chris, and William Bechtel. 2006. "Symbolic versus Subsymbolic" in *Encyclopedia of Cognitive Science*. John Wiley & Sons, Ltd.

Eliasmith, Chris. 2013. *How to Build a Brain: A Neural Architecture for Biological Cognition*. Oxford University Press.

Feyerabend, Paul. 1975. *Against Method*. New Left Books.

Feynman, Richard. 2017. *The Character of Physical Law*. MIT Press.

Fieguth, Paul. 2017. *An Introduction to Complex Systems: Society, Ecology, and Nonlinear Dynamics*. Springer.

Fodor, Jerry. 1998. *Concepts: Where Cognitive Science Went Wrong*. Clarendon Press.

Frege, Gottlob. 1997. *The Frege Reader*. Editor: Beaney, Michael. United Kingdom: Wiley.

Glennan, Stuart. 2017. *The New Mechanical Philosophy*. Oxford University Press.

Gonzales, L. 2004. *Deep survival: who lives, who dies, and why : true stories of miraculous endurance and sudden death.* 1st Norton pbk. ed. New York: W.W. Norton & Co.

Hanson, Norwood. 1958. *Patterns of Discovery*. Cambridge University Press. 30.

Haslanger, Sally. 2000. "Gender and Race: (What) Are They? (What) Do We Want Them To Be?" in *Noûs*. 34(1): 31–55

Haslanger, Sally. 2019. *Conceptual Amelioration: Going on, Not in the Same Way*. Carl G. Hempel Lectures. Massachusetts Institute of Technology.

Hintikka, Jaakko. 2007. *Socratic Epistemology: Explorations of Knowledge-Seeking by Questioning*. Cambridge University Press.

Hoyningen-Huene, Paul. 1993. *Reconstructing Scientific Revolutions: Thomas Kuhn's Philosophy of Science*. The University of Chicago Press.

Huth, Alexander G., W.A. De Heer, T.L. Griffiths, F.E. Theunissen, and J.L. Gallant. 2016. "Natural Speech Reveals The Semantic Maps That Tile Human Cerebral Cortex" in *Nature*, vol. 532.

Isaac, Manuel. 2020. How to conceptually engineer conceptual engineering?. Inquiry. 1-24. 10.1080/0020174X.2020.1719881.

Jen, Erica. 2005. *Robust Design: A Repertoire of Biological, Ecological, and Engineering Case Studies*. Oxford University Press

Kant, Immanuel. 1996. *Critique of Pure Reason*. Trans. and ed. Werner Pluhar. Hackett Publishing Company.

Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. 2012. “ImageNet Classification with Deep Convolutional Neural Networks” in *Neural Information Processing Systems*. 25. 10.1145/3065386.

Kuhn, Thomas. 1962. *The Structure of Scientific Revolutions*. The University of Chicago Press.

Kuhn, Thomas, 1990. “The Road Since Structure” in *The Road Since Structure*, Edited by Conant, James, Haugland, John. 104.

Ladyman, James, James Lambert, and Karoline Wiesner. 2011. “What is a Complex System?” in the *European Journal for Philosophy of Science*. 3 (1): 33–67.

Lissack, Michael, and Abraham Graber. 2014. *Modes of Explanation: Affordances for Action and Prediction*. Palgrave Macmillan. 109–113

Lombrozo, Tania, Joshua Knobe, and Shaun Nichols. 2018. *Oxford Studies in Experimental Philosophy Volume Two*. Oxford University Press.

Machery, Edouard. 2009. *Doing Without Concepts*. Oxford University Press.

Machery, Edouard. 2016. “The amodal brain and the offloading hypothesis” in *Psychonomic Bulletin & Review*.

Margolis, Eric, and Stephen Laurence. 1999. *Concepts: Core Readings*. MIT Press.

Mason, Robert A., and Marcel Adam Just. 2016. “Neural Representations of Physics Concepts” in *Psychological Science 27* (6):904–913.doi:10.1177/0956797616641941.

Massimo, Pigliucci, and Maarten Boudry. 2013. Philosophy of Pseudoscience: Reconsidering the Demarcation Problem. The University of Chicago Press.

Messier, Christian, Klaus Puettmann, and David Coates. 2014. *Managing Forests as Complex Adaptive Systems*. New York: Routledge.

Millikan, Ruth. 2017. *Beyond Concepts: Unicepts, Language, and Natural Information*. Oxford University Press.

Mizumoto, Masaharu, Stephen Stich, and Eric McCready. 2018. *Epistemology for the Rest of the World*. Oxford University Press.

Mobus, George, Kalton, Michael. 2015. *Principles of Systems Science*. Springer.

Murphy, Gregory L., 2004. *The Big Book of Concepts*. Cambridge, MA: Bradford Books.

Nietzsche, Friedrich Wilhelm. 2003. *The Twilight of the Idols and the Anti-Christ or How to Philosophize with a Hammer*. Trans. and ed. Richard Wagner. Penguin.

Page, Scott. 2011. *Diversity and Complexity*. Princeton University Press. 183–195.

Pecher D, Zeelenberg R, Barsalou LW. 2003. "Verifying different-modality properties for concepts produces switching costs." *Psychol Sci*.;14(2):119-124. DOI:10.1111/1467-9280.t01-1-01429

Popper, Karl. 1963. *Conjectures and Refutations: The Growth of Scientific Knowledge*. Routledge & Kegan, Paul.

Prinz, Jesse. 2002. *Furnishing the Mind*. The MIT Press.

Ramsey, William, Steven Stich, and David E. Rumelhart. 1991. *Philosophy and Connectionist Theory*. New Jersey: Lawrence Erlbaum Associates, INC.

Rumelhart, David E., G.E. Hinton, and R.J. Williams. "Learning Internal Representations by Error Propagation." 1985, doi:10.21236/ada164453.

Scharp, Kevin. 2013. *Replacing Truth*. Oxford University Press.

Simon, Herbert. 1962. *The Architecture of Complexity* in *Proceedings of the American Philosophical Society*, Vol. 106, No 6. 468–482.

Simon, Herbert. 1996. *The Sciences of the Artificial*. MIT Press.

Stanford Encyclopedia of Philosophy. Stanford University. Accessed August 27, 2019. https://plato.stanford.edu/entries/kant-judgment/supplement1.html.

Strogatz, Steven. 2004. *Sync: The Emerging Science of Spontaneous Order*. New York: Penguin Books.

Sytsma, Justin, and Wesley Buckwalte. 2016. *A Companion to Experimental Philosophy*. Wiley Blackwell.

Thagard, Paul. 2012. "Procedural Creativity Invents New Methods." Psychology Today. Available at: https://www.psychologytoday.com/us/blog/hot-thought/201208/procedural-creativity-invents-new-methods.

Thagard, Paul. 2014. *The Cognitive Science of Science: Explanation, Discovery, and Conceptual Change*. The MIT Press. 107–140, 109.

Thagard, Paul. 2019. *Brain-Mind: From Neurons to Consciousness and Creativity (Treatise on Mind and Society)*. Oxford University Press.

Terrence, Stewart, Choo, Feng-Xuan; Eliasmith, Chris. 2014. “Sentence processing in spiking neurons: A biologically plausible left-corner parser.” Proceedings of the Annual Meeting of the Cognitive Science Society. Volume 36.

Weisberg, Michael. 2013. *Simulation and Similarity: Using Models to Understand The World*. Oxford University Press. 156–170, 98–113.

Werning, Markus, Wolfram Hinzen, and Edouard Machery. 2012. *The Oxford Handbook of Compositionality*. Oxford University Press

Wimsatt, William C., 2007. *Re-Engineering Philosophy for Limited Beings: Piecewise Approximations to Reality*. Harvard University Press. 196–197, 358, 141–145.