Conceptual Engineering in the Biological Image

Brian Kobylarz*

bwkobylarz@gmail.com

Abstract

Conceptual Engineering, and investigations into concepts generally, is becoming a focus of recent academic philosophy. The importance of this movement is upheld here, and I approach the conversation with a metaphysics of concept relations. I claim such relations between concepts feature properties expressible in the language of theoretical biology, particularly as presented by William Wimsatt. By synthesising the properties of robustness and entrenchment with themes in the conceptual engineering conversation, I introduce new concepts of my own and hope to show that a biology, or complexity orientation, can aid the conceptual engineer.

1 Introduction

The creation of concepts is an activity at least as old as philosophy itself, and meta-level considerations of this activity can be found throughout the history of philosophy. Recently, convergence has emerged on the importance of this meta-level, and attempts are being made at teasing out the various dimensions to concepts and their production. There exist, for instance, conceptual engineers and meta-philosophers who advocate the value of concept building and selection generally, with implications ranging from the epistemic to the normative, and with import into our day-to-day (Burgess and Plunkett 2013; Eklund 2017). This effectively amounts to a methodological shift from using whatever concepts cultural evolution has left lying around to the intelligent design of concepts. Another group, the cognitive penetrability community, suggests an entanglement between our conceptual apparatuses and our ability to perceive (Zeimbekis and Raftopoulos 2015). Here we find a successor to the theory-ladenness of observation tradition in the philosophy of science (Hanson 1958; Kuhn 1962; Thagard 1992). Psychologists and neuroscientists, through various brain imaging technologies, are exploring concepts on empirical grounds (Binder et al. 2009; Hofmann et al. 2012). Recently, a study at Carnegie Mellon

*Brian Kobylarz completed his BA in Philosophy at William Paterson University, and plans on submitting his transcripts to PhD departments in 2019. He is interested in various philosophical traditions but is currently most interested in concepts, how they grant agency, and how to model it. He is currently exploring intersections between Conceptual Engineering and Dynamical Epistemic Logic.

claims to have visually confirmed distributed neural activation signatures which corresponded to subjects thinking about particular concepts in physics (Mason and Just 2016). Along the meta-philosophical axis, Luciano Floridi presents a call-to-action of his own:

> Too much ink has been spilt on philosophy as conceptual analysis. The alternative view, that philosophy is at least as much, if not actually more, engaged with creating, refining and fitting together our conceptual artefacts [...] has received too little attention (Floridi 2011, 293).

A trans-disciplinary convergence exists here in terms of the significance granted to concepts. I too share many of these beliefs, and I am convinced further convergence is possible, desirable, and important. Concepts are given many roles, they are said to frame discourse, to grant agency, orient observation, compose into explanatory frameworks, and allow for reasoning generally. It is difficult to even speak of 'concepts' without employing the usage of the word 'concept.' As with any subject matter imbued with such gravity there exist less optimistic, sceptical voices. One such representative here, expressing important issues for those who wish to engineer concepts, is Herman Cappelen. For Cappelen our concepts—even though he refuses to label them as such—are largely at the whim of social forces; they are not things we can domesticate and sterilise nor intelligently design on par with what the conceptual engineer hopes. In his own words:

> In most cases the detailed mechanisms that underpin particular instances of conceptual engineering are too complex, messy, nonsystematic, amorphous, and unstable for us to fully grasp or understand (Cappelen 2018, 72).

I find Cappelen's portrayal to be the right portrayal, but I differ in that I believe a solution rather than a dead end is implied. I thus find myself on more optimistic grounds. I take the usage of 'complex' to be where I start. By 'complex' I mean a system with many moving parts, something that if modelled, would demand many variables to capture the relevant mechanisms such that correlations and causal relationships can be mapped, explored, and exploited. Often such systems exhibit emergent behaviours, and non-linear developments which are beyond prediction. If conceptual engineering has a complexity problem, then looking towards those sciences which have successfully managed to navigate complexity appears to be the reasonable step to take. This paper thus exists as an initial, and very rudimentary attempt at exploring some points of connection between conceptual engineering and the complexity/network sciences. William Wimsatt happens to be an exemplar in providing novel ways to speak about and understand complexity. As such, I will be appropriating some of Wimsatt's language to the conceptual engineering conversation. I believe this labour to be one which is in some sense continuous with Cappelen as he claims: 'We should object to any theory that makes it [conceptual engineering] easy' (70). By portraying concepts, and their relations in light of complexity, we can admit the messiness while welcoming various techniques and methods for navigating the messiness.

What I intend on importing here exactly, from Wimsatt to conceptual engineering, are two ideas: robustness, and entrenchment. Both ideas denote object relations. The modus operandi of this paper can then be said to be descriptive——a metaphysical investigation into concept relations where I hope to show that concepts relate as such, and that the consideration of both relations are important for the engineering, augmentation, and implementation of concepts. It may seem strange to apply this biological image to concepts, as it appears to demand more metaphysics than I have room for here, but as Wimsatt's work makes explicit: the properties under consideration are found beyond the biological sphere. Generally, a network that adapts and has dependency relations between its members will show the properties I will explore (Wimsatt 2007). Thus, it is equally fair to refer to these *biological properties* as *adaptive network properties* more generally. Once we grant adaptive network properties, it should be less controversial to associate robustness and entrenchment with concepts and their functionality.

2 Divisions of compositional labour

To frame, take for instance the concept 'ice-cold', this may be used to refer to a drink's temperature. Given the right conditions, this concept can also be employed to refer to a particularly cruel individual. For the concept 'ice-cold' to survive, and to be suited for a range of phenomena like the temperament of an individual or the temperature of a drink, it must change along various axis, including those denoting verbal fashion, the environmental context which influences its utterance, and the analogical skills of the speakers in the conversation where 'ice-cold' appears. I find it, then, non-controversial that concepts are adaptive, as their applicability, usage, and utility (in terms of possible inferences, and actions they enable) change relative to various influences.

Another influence on concepts can be concepts themselves. Here we see network or ecological properties present in terms of how our concepts relate to one another. Consider the concept 'fire', which when put into conjunction with the concept 'danger' can be altered such that it can convey a very particular warning, and with it a sense of urgency that can trigger the production of stress hormones. The same concept 'fire', when conjoined with 'barbecue', can have a welcoming effect such as the management of hunger, leading to inferences that dinner is on its way. Because of such relations between concepts and how they can influence one another, I then have no issue committing myself to what can be called *concept ecologies* or *concept networks*. Dependency relations between the concepts can be elucidated through this language. We use concepts, and they can loosely be said to mutate by forces both internal and external. The internal amounts to concepts operating on concepts, such as how they compose to enact different responses as when 'fire' is used in conjunction with 'danger'. The external can denote those environmental forces, such as an 'ice-cold' person, which influence what utterances are made (which concepts are enacted). I accept both divisions of labour. While we may have our differences as to how we weigh the internal labour and the external labour which shape our concepts, I find concept ecologies or concept networks to be permissible for both internalists and externalists, as well as those who uphold the necessity of the conjunction of both, such as myself. This essay will, however, be largely focused on the internal relations, in other words: the concept-on-concept relations within a concept ecology. It is my belief, then, that for the kind of mastery over concepts the conceptual engineer desires, the complexity of the related systems must be considered. I will begin by outlining robustness, move to entrenchment, and conclude with some remarks as to what other problems arise within this framework.

3 Robustness and amelioration

Wimsatt defines robustness as 'the existence of multiple means of determination or access to an entity, property, process, result, or theorem that are at least partially independent' (359).

In biology, this lends itself well to expressing situations where one biological activity can be instantiated by various biological mechanisms. The imagery of networks can also be grasped through Wimsatt's characterisation. Here, nodes can represent an entity, and arrows between the nodes can represent the causal relationships between nodes in a network. The more arrows which converge upon a single node (entity), the higher the measure of robustness, the greater chances an entity can be instantiated by other entities within its ecology (network). The evolutionary benefit here is obvious, as mechanisms can pick up the slack of other mechanisms—a network feature which increases the survivability of an organism.

Let's take concepts then as our entities. Robustness can then be held as the measure which allows us to approximate the degree of effect concepts on other concepts, and what effects a network can produce in light of this relation. The concept designer, nearly by definition, is someone who is interested in the result(s) of a given concept. Take for instance a general situation where X is a desired effect believed to be arrived at through concepts 2, 3, and 4. In our concept ecology, concepts 2, 3, and 4 are present, but none of them are producing X. If X is not currently actualised by this network, then it could amount to an unnecessary and costly endeavour for the conceptual designer who chooses to build an entirely new concept to bring about X. The payoff may be arrived at through investing less into other—already existing—concepts by way of a kind of augmentation to them. Choosing to build a new concept here would be akin to nature building an appendage where one isn't needed, as the present appendages would be able, with slight modification, to complete the desired task. To understand robustness is then to both understand a property of concept ecologies, and to know when concept creation is appropriate—and when augmentation of already existing concepts is appropriate.

To make this less abstract, let's return to the concept 'fire' where we already found an example of robustness. Someone possessing the concept 'fire' has the initial materials which enables inferences to be arrived at for a robust range of situations where 'danger' and 'comfort' are appropriate. If their 'fire' concept is defective, lacking a degree of robustness in terms of possible inferences regarding the dangers of fire, then we may opt to educate them in terms of the dangers of fire, and fire safety more generally, thus expanding the possible inferences and actions the concept 'fire' can grant a subject. In the language of conceptual engineering, *amelioration* is the name given to the activity of augmenting concepts as such. As Herman Cappelen understands, amelioration can come in a 'variety' of strategies, one such example Cappelen articulates:

> You have a word with a certain meaning and extension/intension. You keep the word and improve the meaning/extension/intension. In the standard case, the lexical item will be preserved, but the semantic value will be improved (Cappelen 2018, 35).

In the case of teaching fire safety, we find the lexical item preserved ('fire'), and the semantic value changed such that 'fire' denotes not only that which cooks, heats, and warms but also that which can harm, kill, and destroy. Teaching fire safety, in other words, makes for a more robust conception of 'fire', broadening the range of possible inferences regarding fire and thus the possible actions someone sufficiently informed can take in relation to fire. One can imagine such an expansion to allow other effects to be arrived at through combinatorial concept processes such as when 'fire' combines with 'danger' in a statement, allowing both concepts to interface with one another to grant the desired effect of successfully conveying a warning.

In closing this section, there are those who find amelioration to be too difficult. I find this position odd as it seems clear to me that while amelioration is difficult, there exists examples of it. My initial intuition is to consider education generally and how at all it would be possible if amelioration were beyond us. Propaganda also comes to mind, where it can be argued that significant alterations have been made throughout history to change the extensions of concepts such as 'enemy', 'solidarity', 'terrorism' and 'terrorist', all of which appear malleable in the hands of propagandists. While this amounts to a generally monstrous conceptual engineering, it also suggests a more optimistic reading: that concepts are within our control and that it can presumably be employed for good. A third example comes from psychology, where we find that cognitive behavioural therapy has for decades demonstrated success for the treatment of various mental illnesses (Hofmann et al. 2012). The patient, employing cognitive behavioural therapy, is directed to keep journals, to stay mindful of negative thinking habits, and to replace those negative thoughts with more positive conceptions. This amounts partially to a reordering of already existing cognitive architecture, and one which is conducted with success by an individual without the funding and mobilisation of propaganda and education. I believe one can frame conceptual engineering as a kind of cognitive behavioural therapy for epistemic and normative conepts. Amelioration, as far as I can tell, is a ubiquitous human activity.

4 Generative Entrenchment

As William Wimsatt presents, generative entrenchment is 'a measure of how many things depend upon an element and thus likely to change if it changes' (Wimsatt 2007, 355).

We can think of our anatomical arrangements as deeply entrenched—they depend upon a number of mechanisms such as our genetics. This amounts to a dependency between a higher order thing and a more foundational thing where a change to our genetics can have significant ramifications to our anatomical arrangement. Entrenchment can also play a role in determining what entities can be integrated into an ecology offering a degree of resilience, and predisposition, to a network where whatever is commensurate with higher order entities must also be commensurate with lower level entities due to the dependency relations. In other words, entrenched entities operate as gatekeepers, in part determining what is integrated into our concept ecologies, and what is not.

As I mentioned earlier, concepts appear to depend on other concepts too, as does our agency itself. This important point is expressed by Burgess and Plunkett in their work 'Conceptual Ethics I':

> Arguably, our conceptual repertoire determines not only what beliefs we can have but also what hypotheses we can entertain, what desires we can form, what plans we can make on the basis of such mental states, and accordingly constrains what we can hope to accomplish in the world. Representation enables action, from the most sophisticated scientific research, to the most mundane household task. It influences our options within social/political institutions and even helps determine what kinds of people we can be, what sorts of lives we can lead (Burgess and Plunkett 2013, 1096–97).

We are presented with concepts as influencing one another such that our epistemic capacities are implicated by the concepts we possess, as are our capacities to act more generally. As the conjunction of entrenchment and concepts suggest here: to what degree are our concepts then entrenched by one another, and ultimately what does this mean for the constraint and freedom we have to know and act? The ability to see concept ecologies as such—to conceptualise concept ecologies as such—leads to investigations into some of the necessary conditions for concept implementation. The concept designer speculates in the form of: if concept 1 is to be introduced into concept ecology ϕ, then what entrenched concepts are present in concept ecology ϕ such that concept 1 may be successfully or unsuccessfully integrated? An example: One could imagine, for instance, a fairly reasonable individual with entrenched concepts which make said individual value virtues such as deduction and consistency. This individual's concept ecology may then be immunised against the addition of new concepts which do not follow, either explicitly or implicitly from the concepts present. In other words, our entrenched concepts can immunise us against some concepts, and leave us predisposed to others. If Burgess and Plunkett are correct, this greatly determines what we are capable of. Conceiving of concept relations with entrenchment can aid the concept designer in terms of what concepts will survive and what concepts will die off when integrated into a concept ecology.

It is common in biology to find mechanisms and processes which have multiple functions. While entrenched concepts can immunise against or predispose us to new concepts, they can also have large scale implications when tampered with, given the nature of their fundamentality. I am forced to consider, then, the possibility of *concept ecology collapse*. Would it be possible, once the entrenched concepts are discovered, to manipulate them or introduce concepts which cause the networks to fail? Are there concepts which entrench to such a degree that entire branches in our *concept networks* would fail if the entrenched concepts were tampered with in an irresponsible way? The implications for this are as exciting as they are horrific. It may be useful to borrow another concept from biology here, that of the 'invasive species', where a species, traditionally foreign to an ecology, can be introduced with disastrous consequences, due to an ecology lacking an arrangement which can support such a species. The conceptual designer may want to consider *invasive concepts*, or the degree to which a designed concept is invasive and how such invasive concepts function relative to a concept ecology and the desired goal(s).

Perhaps here amelioration can find itself in some cases to be a supplement to concept creation and implementation where an amelioration is conducted on deeply entrenched concepts while new concepts are introduced, creating an environment where a new concept can be integrated with greater ease while diminishing potential externalities including those such as concept ecology collapse, or the rejection of a new concept given the entrenched concepts inherent to a concept ecology. Entrenchment allows further evaluations of the structure of a concept ecology, giving the concept engineer an idea of what the fundamental concepts are, and how to proceed.

5 Conclusion and future work

Experiments conducted in the 1980s lead to the understanding of fruit fly construction during the embryonic stage where the development of a fruit fly emerges from head to wings. This process follows an elaborate and delicate chemical activation of certain genes in a particular order (Shubin 2009). The activation is sequential, and thus the history of this activation is necessary to fully grasp and model the development of a fruit fly. . There exist numerous biological mechanisms throughout nature which exhibit sequential operations. Delicate operations, such as this, share some symmetry with conceptual engineering. For mastery over our concepts, precision seems necessary, and precision appears only possible when the multivariate and complex dynamics which shape our concepts are understood. This amounts to an advocacy of descriptive investigations into our concept-and-concept relations as well as the concept-and-world relations which determine our ecologies/networks. This of course demands much more work than what is possible in this paper, which as I mentioned in the introduction is an attempt to simply bridge complexity and conceptual engineering in an initially novel way.

When one looks towards biology and other sciences, which take seriously the hierarchical and complex relations between their objects, there exists measurements. Robustness and entrenchment can be defined quantitatively, if my characterisations have been successful then presumably measurement can be on the horizon for the conceptual engineer. Significant work is needed in order to arrive at such a methodology. This leads to the elephant in the room, that being what exactly a concept is, as before we can begin to model, map, measure and determine with empirical accuracy our concept ecologies/networks, we must have a satisfying definition of what is to be measured. Cappelen's cut is deeper here as he addresses the lack of convergence in terms of a universally accepted theory of concepts (Cappelen 2018). This attack on the presuppositions of conceptual engineering may condemn the project before it gets off the ground. It is true that there exists a smorgasbord of different theories of concepts (Carey 2009; Machery 2009; Churchland 2012; Gärdenfors 2000), and that as such a pessimistic reading is within reason. I do not have the room here to develop a theory of concepts, but there are some presuppositions which this paper makes concerning the nature of concepts. One such presupposition is that concepts are representational devices as they approximate environmental conditions. On the brain level, these representations are seemingly distributed (Binder et al. 2009). Another presupposition present is that concepts have a combinatorial nature, they arrange themselves relative to one another and environmental context. I find that the compositionality of our concepts constrains and determines much of what our concepts can do. As such, this biology/complexity semantics I have explored seems necessary to consider beyond the neural-level. This amounts to a descriptive venture into a level beyond neurons (concepts and their relations) and the prescriptive claim that a science of such a level is necessary. Thus, I am at odds with the Eliminativist tradition. Both the significance of compositionality (see Werning, Hinzen and Machery 2012; Murphy 2004) and concepts as representational devices (Churchland 2012; Mason and Just 2016) are popular theories of concepts, and their performance. How both levels interface with one another is another problem. I suspect that as neuro-imaging technologies advance, concepts themselves will begin to be taken more and more seriously as observation of conceptual operations in the brain will come into clearer focus. The work I've presented here, I believe to be useful to the understanding of such an interface between neural and language based investigations/interventions into our concepts. While I am interested in the emancipatory capabilities of conceptual engineering, I then am also interested in a kind of concept-informatics which accompanies the engineering of our concepts.

I believe both would feedback into one another and that there are answers, with number values, to questions such as:

(1) How deeply entrenched is concept 1?

(2) What degree of robustness, in terms of possible inferences, does concept 2 possess?

Finally, I am of the position that conceptual engineering needs to be an interdisciplinary venture. I believe that organisational attempts to synthesise cross-discipline discoveries should be made, and that each discipline contributes to a greater image of concepts and their relations, which hang together as complex systems. The philosopher's labour enters in terms of the creation of concepts to grant greater degrees of agency, and the augmentation of concepts to make actual the unfamiliar potentials in familiar concepts. This extends to the necessary meta-labour needed to sort out conceptual engineering itself, which I have hoped to elaborate a bit further myself.

References

Binder, Jeffrey R., Rutvik H. Desai, William W. Graves and Lisa L. Conant. 2009. 'Where Is the Semantic System? A Critical Review and Meta-Analysis of 120 Functional Neuroimaging Studies'. *Cerebral Cortex* 19 (12): 2767–96. doi:`10.1093/cercor/bhp055`.

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

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

Carey, Susan. 2009. *The Origin of Concepts.* Oxford University Press.

Churchland, Paul M. 2012. *Plato's Camera: How the Physical Brain Captures a Landscape of Abstract Universals.* MIT Press.

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

Floridi, Luciano. 2011. 'A Defence of Constructionism: Philosophy as Conceptual Engineering'. *Metaphilosophy* 42 (3): 282–304.

Gärdenfors, Peter. 2000. *Conceptual Spaces: The Geometry of Thought.* MIT Press.

Hanson, Norwood Russell. 1958. *Patterns of Discovery: An Inquiry Into the Conceptual Foundations of Science.* Cambridge University Press.

Hofmann, Stefan G., Anu Asnaani, Imke J. J. Vonk, Alice T. Sawyer and Angela Fang. 2012. 'The Efficacy of Cognitive Behavioral Therapy: A Review of Meta-analyses'. *Cognitive Therapy and Research* 36 (5): 427–40. doi:`10.1007/s10608-012-9476-1`.

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

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

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

Murphy, Gregory L. 2004. *The Big Book of Concepts.* MIT Press.

Shubin, Neil. 2009. *Your Inner Fish: The amazing discovery of our 375-million-year-old ancestor.* Penguin.

Thagard, Paul. 1992. *Conceptual Revolutions.* Princeton University Press.

Werning, Markus, Wolfram Hinzen and Edouard Machery, eds. 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.

Zeimbekis, John, and Athanassios Raftopoulos, eds. 2015. *The Cognitive Penetrability of Perception: New Philosophical Perspectives.* Oxford University Press.