The architecture of higher thought

Daniel A. Weiskopf

1. The hierarchy of thought

The idea that mental activity can be arranged according to a hierarchy of complexity has ancient origins, with roots in the Aristotelian psychological division of the soul into nutritive, perceptual, and intellectual faculties. Elementary functions dedicated to sustaining life undergird those that engage cognitively with the sensible world, which in turn support the active faculties of abstract reason. Higher faculties are meant to be those that have, in some sense, greater abstraction, generality, or complexity, which links them ethologically with distinctively human cognitive capacities, and developmentally with the mature state of the human cognitive system as opposed to that of the infant or child.[1]

This distinction recurs at the inception of modern scientific psychology in Wundt, who sharply distinguished between sensory and perceptual capacities and other mental capacities such as learning and memory. For Wundt, the distinction between the two was not just between two types of cognitive process, but also implied a methodological dualism: sensory capacities were grounded in species-typical physiological mechanisms, while higher thought, insofar as it is conditioned by language and culture, must be studied using the methods of sociology and anthropology, as part of a distinctive *Völkerpsychologie* (Hatfield, 1997). This separation continued well into the 20th century. Neisser (1967), for instance, does not share Wundt's methodological dualism, but his landmark textbook on cognitive psychology is divided into two major halves, one dealing with visual processing and the other with auditory processing, with higher cognitive processes (memory, language, and thought) confined to the brief epilogue.

[1] For discussion of the Aristotelean divisions and related ancient notions, see Danziger (1997, pp. 21-35).

In this paper I critically assess existing accounts of what it means for there to be a hierarchy of thought, and offer an alternative taxonomy that views higher cognitive capacities, paradigmatically exemplified by conceptualized thought, as instantiating three distinct but convergent functional profiles. This analysis can overcome problems that plague other accounts, which tend to be hopelessly underspecified or oversimplified; moreover, it shows an important sense in which the apparently disunified higher faculties can be seen to belong to genuine psychological kinds.

2. Attempts to analyze the hierarchy

The simplest account of the higher/lower division follows Wundt, Neisser, and much of the rest of the psychological tradition in drawing a sharp line at the senses. On this view, lower cognition is made up of perception and action systems, with higher cognitive processes being the remainder, including such things as language comprehension and production, memory encoding and access, decision-making, categorization, planning, analogy and creative thought, and various forms of reasoning: deductive, inductive, and abductive. Sometimes added to the list are forms of metacognitive awareness, both of one's own thoughts and those of others, as well as metacognitive control over one's thought processes. Other cognitive states and systems such as the emotions and other affective, motivational, and drive systems are harder to place on this account, although the emotions have widely noted affinities with perceptual states.

An immediate problem with this list is that it fails to tell us what higher processes have in common, aside from merely *not* being sensorimotor: the higher processes are represented as a motley and disunified lot. *Prima facie*, though, we should attempt to seek out and display theoretically unifying features in cognitive systems wherever they occur. Call this the *Unity Problem*.

A further problem is that as it stands the list fails to make enough distinctions. There may be a variety of processes that are not sensorimotor but which are not intuitively higher processes. For example, elementary grasp of number and quantity seems to involve a rapidly activated analog magnitude system that estimates the number of distinct items in an array, and which can distinguish two separate arrays as long as they stand in a favorable ratio (roughly, obeying Weber's law). But this system is amodal, or at least supramodal, and hence is not part of any particular perceptual system. At the same time, the numerical estimations and computations it can carry out are highly limited (Beck, forthcoming), falling far short of full-fledged adult numerical competence. The mind appears well-stocked with intermediate systems of this type (Carey, 2009), which are not captured by a simple binary distinction. Call this the *Richness Problem*: we need our taxonomy to make *enough* distinctions for all the systems we find.

More worrisome is that sensorimotor systems themselves may function by making use of what appear to be some of these very same higher processes (Rock, 1983). Consider categorization. There is no clear and unambiguous notion of what a categorization process is, except that it either assigns an individual to a category ("a is F") or makes a distinction between two or more individuals or categories ("these Fs are G, those Fs are H"). But many cognitive systems do this. Vision is composed of a hierarchy of categorization devices for responding to edges, textures, motion, and whole objects of various sorts, and the language faculty categorizes incoming sounds as having various abstract phrasal boundaries, determines the phonetic, syntactic, and semantic classes that words belong to, and so on. Indeed, almost every cognitive process can be seen as involving categorization in this sense (van Gelder, 1993).

Similarly for reasoning. It is widely assumed that perceptual systems carry out complex inferences formally equivalent to reasoning from the evidence given at their inputs to the conclusions that they produce as output, as in models that depict visual analysis as a process involving Bayesian updating (Knill & Richards, 1996). Language comprehension systems must recover the structure of a sentence from fragmentary bits of evidence, a task that has essentially the form of a miniature abductive inference problem. So the mere existence of a certain type of process cannot draw the needed distinction, since these processes may occur in both higher and lower forms.[2] Call this the *Collapse Problem*: the difference between higher and lower faculties must be a stable and principled one.

A related way of drawing the higher/lower distinction says that lower cognition takes place within *modular* systems, while higher cognition is non-modular. This too requires more precision, since there is a range of different notions of modularity one might adopt (Coltheart, 1999). In Fodor's original (1983) discussion of modularity, the line between modular and nonmodular systems coincides roughly with the line between higher and lower systems: modular systems are, *de facto*, the peripheral input and output systems (plus language), while central cognition, the home of the folk psychologically individuated propositional attitudes, is the paradigm of higher cognition. But unlike on the simple account, there can be a whole sequence of modular processes that take place before arriving at central cognition.

A virtue of using modularity to distinguish higher and lower systems is that nonmodular systems can be given an independent positive characterization. The signature phenomena associated with central cognition are the properties of being Quinean and isotropic. Quinean processes are those that are sensitive to properties of a whole belief system, while isotropic

---

[2] This point is made by Stich (1978), who attempts to distinguish doxastic states and processes from subdoxastic ones on the grounds that the former are inferentially integrated and conscious. Inferential integration will appear as one of the criteria for higher cognition on the account developed here, although not consciousness.

processes are those that in principle may access any belief in the course of their operation. These two properties are, in turn, best explained by central cognition being a domain-general and informationally unencapsulated system. The ascent towards higher cognitive processing involves the progressive loosening of the constraints of modularity: greater generality in processing and wider access to stored information. This corresponds to the classical understanding of canonical higher thought processes such as reasoning, analogy making, and decision processes. There is thus some account of how to recognize whether a process belongs to higher or lower cognition, unlike on the simple view on which higher processes are more or less just those that are not perception.

However, the necessity of these conditions is challenged by massively modular models of cognition. In these architectures, even central cognition is fragmented into a host of domain-specific processors. Moreover, none of these systems has the unrestricted scope and access characteristic of Fodorian central cognition. In the standard formulation of evolutionary psychology, cognitive domains corresponding to adaptive problems are assigned to dedicated reasoning systems. Mate selection, predator detection, conspecific recognition, cheater detection, and a host of other survival-linked cognitive skills all employ distinctive reasoning principles, tap separate bodies of information, and operate semi-autonomously from each other. The total interaction of all of these systems produces the rich complexity of human thought, rather than any single domain-general reasoning device (Carruthers, 2006).

At issue here is not whether the mind is in fact massively modular; the point is that as far as constructing a taxonomy of higher and lower systems goes, modularity by itself does not seem to capture what we are looking for, since it seems perfectly coherent for there to be modularly organized higher cognitive systems. Call this the *Neutrality Problem*: in attempting to reconstruct the distinction, we should try to do so in a way that accommodates massively modular architectures (and others) in principle.

Finally, consider two non-psychological accounts of the higher/lower distinction. One is drawn from neuroscience. The brain offers what seems to be an obvious model of higher and lower systems, namely one grounded in afference hierarchies. Moving inwards from the peripheral nervous system, each synaptic step represents a rung on the ladder towards ever-higher brain regions. This model is implicit in theories that assumed that brain network connectivity has a straightforward hierarchical structure, with inputs flowing to a central dominant control point before turning into efferent signals and flowing outward again. Activity is propagated linearly upwards through sensory areas, through "association cortices", and ultimately to wherever this seat of control is located—perhaps the distinctively human prefrontal cortex.[3]

This model offers a literal interpretation of our central metaphor: higher cognitive functions are localized at higher levels of the afference hierarchy. The model also gains support from the fact that sensory systems typically involve internal processing hierarchies in which they move from simpler to more abstract and complex representations of their objects: from edges and various other primitive features to 2-D shapes, and from shapes of objects to assignments of category membership. Higher cognition involves extracting and manipulating such abstractions. Since they are the peak of the afference hierarchy, higher brain regions in this scheme would also receive and integrate inputs from all lower regions, which puts them in a position to implement informationally unconstrained reasoning and decision-making processes. On representational and

[3] So Benton (1991) comments that historically the prefrontal region was thought to provide "the neural substrate of complex mental processes such as abstract reasoning, foresight, planning capacity, self-awareness, empathy, and the elaboration and modulation of emotional reactions" (p. 3). For the history of ideas about "association cortex" more generally, and about the localization of intellectual function, see Finger (1994), Chs. 21 & 22.

control-theoretic grounds this might seem to capture some important properties of higher thought.

Unfortunately, the picture of neural organization that this model relies on is false when examined in detail. Within only a few synaptic hops inward from the periphery, it rapidly becomes impossible to trace any single neat hierarchy of regions. While afferent pathways are initially organized into linear tracts, widespread lateral branching structure emerges early on. Even in primary sensory areas there is massive crossover; early sensory processing rapidly becomes multimodal, and intertwined pathways are the norm (Ghazanfar & Schroeder, 2006; Schroeder & Foxe, 2004). Many of these pathways are also recurrent, forming loops rather than unidirectional paths. Such top-down influences are essential for resolving ambiguities in the perceptual array and speeding recognition (Kveraga, Ghuman, & Bar, 2007). Finally, there is no single neural pinnacle where these tracts converge. Higher cognitive processes recruit widespread networks activating multiple regions at many putatively distinct levels of processing (Anderson, 2010). Reading higher processes off of the neural architecture, then, seems unpromising.

A final non-psychological account appeals to evolutionary history. Higher cognitive processes on this view would be those that are of more recent evolutionary origin, with lower processes being earlier and (therefore) simpler, and higher processes being those that build on these and enable more complex and powerful cognitive achievements. This notion goes back to T. H. Huxley and his "doctrine of continuity", along with the work of C. Lloyd Morgan on the varieties of animal intelligence, and ultimately to Darwin himself, who in *The Descent of Man* posited "numberless gradations" in the mental powers of nonhuman animals.[4]

[4] Danziger (1997, pp. 66-84) provides an excellent survey of the emergence of the modern term "intelligence" and the debates concerning its proper application to animal cognition and behavior.

However, this proposal faces several objections. For one thing, even structures that make their first appearance early in a species' history are subject to continual modification. This is true even for the visual system, which shares many commonalities across existing primate lineages; human area V1 exhibits distinctive specializations in laminar organization and neuronal phenotypes and spacing, and area V3A exhibits greater motion-sensitivity than does its macaque homologue (Preuss, 2011). If it is the current species-typical form of the capacity that counts, these too may have a relatively short evolutionary history. As a species undergoes new phenotypic modifications, these systems may adjust as well; as new capacities are added, they modify the old ones in various ways. The addition of language areas to the brain, for example, does not leave the underlying architecture the same, with widespread changes in connectivity within and among areas that have existing primate homologues (Barrett, forthcoming). Even more surprisingly, bipedal walking seems to recruit areas in and around primary visual cortex independently of whether visual input is being processed (Allen, 2009, p. 89). So there may be no stable temporal order of fixed systems to appeal to; what these systems *do* changes as new components are added around them.

Moreover, some systems that do not fit the intuitive profile of higher cognition may be comparatively recent, and some higher cognitive structures may be ancient. According to some computational models, emergence of language may depend on the development of an expanded short term memory buffer for temporal sequences, but this expanded buffer *per se* does not seem to be an especially "higher" capacity (Elman, 1993). For an example of the latter, Carruthers (2004) has argued that even bees and other insects may literally have beliefs and desires, making them positively ancient in evolutionary terms. Whether this is true or not (see Camp, 2009 for criticism), it doesn't seem that there is any *essential* link between being on the list of intuitively higher cognitive states and falling in any particular place on the evolutionary calendar.[5]

3. The functional components of higher cognition

Many of the existing accounts of how cognitive faculties can be ordered and distinguished, then, are non-starters when examined in detail. But considering their defects may point the way towards a more adequate account. For present purposes, I will take conceptualized thought as the paradigmatic form of higher cognition. If we can get clear on the functional role that concepts play in cognition, we will have a grasp of what at least a large chunk of higher thought consists in. We can then proceed to see whether the notion may be generalized in any way. In defense of this strategy, I take it that concept possession is special—not any cognitive system automatically counts as having concepts. Mature, normal humans have concepts, as do very young humans that are developing normally; and some animals may, although this is an open question. The usefulness of having a general criterion for concept possession is precisely that it allows us to make these sorts of developmental and ethological distinctions.

I suggest that there are three properties that characterize the conceptual system. These properties are logically and causally independent, but they collectively describe what concept possessors share. They are:

1. Representational abstraction
2. Causal autonomy
3. Free recombination

[5] For a reconstruction of the phylogeny of human cognition that depicts it as emerging from a graded series of broadly holistic modifications to the brain occurring at many levels simultaneously, see Sherwood, Subiaul, & Zawidzki (2008).

These are not "analytic" claims or stipulative definitions, nor do they necessarily express what we mean in everyday life by using the term "concept". These are intended as part of a theoretical definition aimed at carving out a functional and explanatory role that distinguishes concepts as such from other kinds posited in our psychological theories.[6]

3.1. Representational abstraction

Take *representational abstraction* first. Concepts are mental representations, but there are many types of mental representations besides concepts. There are perceptual representations in various sense modalities and submodalities, and there are motor representations that control how we are to move our bodies. There are representations of quantity and number, at least some of which appear to be amodal or cross-modal. There are various spatial, temporal, and geometric representations; for example, representations of where we are located in space, where we have traveled recently, where landmarks in the environment are, and so on. And language users have a host of representations of the phonological, morphological, syntactic, and semantic properties of words and sentences.

Concepts, like other representations, are individuated in part by what they represent. A concept is always a concept *of* a certain entity (my cat, the man in the silver jacket, the Eiffel tower, the number 3), property (being furry, being a third son, being a quadruped, being irascible), kind (cats, unicorns, gluons, Hilbert spaces), state (being true, being dead, feeling elated), or event (exploding, kicking, kissing). Some of the things that we can have concepts of are things that can also be represented by us in other ways. We can perceive red, furry things and we can think about redness and furriness. These properties are represented both in perception and conceptually. The property of being a triple may be perceptually detectable and manipulable by

[6] On the picture of functional kinds being employed here, see Weiskopf (2011).

our faculty of numerical cognition, but we can also entertain the concept of the number 3. Similarly, we learn to detect and represent nouns long before we learn the concept of a noun. The mind's systems are representationally redundant in this way.

But many things *cannot* be represented by our dedicated systems for perception, action, numerical cognition, language, etc. For instance, Hilbert spaces, being mathematical objects, have no perceptual manifestations and surely exceed our number faculty's powers. Whether something is an allusion to Dante's *Inferno* is not something that the language faculty is capable of representing, and yet we can, with training, come to recognize such allusions. The same goes for social roles such as the ancient Greek notion of being an *erastes* or *eromenos*, or the contemporary relation of being a "friend with benefits"; for aesthetic qualities like being an example of Minimalist composition or being a sonnet; for material kinds such as polyester-rayon blend fabric; for theoretical kinds in various sciences such as extrasolar planets, ribosomes, and prions; and so on. We can represent and think about all of these types of things, but the capacity to do so does not seem proprietary to any cognitive system, particularly not our sensorimotor systems.

To say that concepts are abstract representations, then, is to say the following. First, we are capable of having concepts of things that we cannot represent using any of our other representational systems. Our concepts can *transcend* the representational resources of these systems. As many of the above examples indicate, this can happen where the categories we are representing are ones that simply have no perceivable or manipulable manifestations at all, and hence cannot be directly represented by sensorimotor systems.

Second, not only can we exceed the representational power of these systems, we can cross-classify the world relative to the distinctions made by these systems. Where perception groups objects together (e.g., by their shape, size, color, motion, and so on), we can distinguish them in conceptualized thought. Two kinds of berry look similar, but are actually distinct; jadeite and nephrite are similar in their perceivable traits, but differ in molecular composition. And where perception groups objects separately, we can bring them under a common concept; hence we recognize that the caterpillar and the butterfly, appearances aside, are one and the same.

Third, we are also at least sometimes capable of having concepts of the very same categories that are represented by these other cognitive systems. However, our concepts need not represent them in the same way as these systems do. Hence there can be concepts of perceivable objects that do not represent them in a perceptual way. This is clear from the fact that we can think about entities without knowing their appearance, and we can perceive their appearances without knowing what they are. Concepts may re-encode categories that are represented elsewhere in the mind, and some of these re-encodings involve representing the same category but discarding some of the information carried by earlier forms of representation (e.g., fine-grained information about appearances).

To summarize these points, we can say that a system S_1 has representational resources that are more abstract than S_2 when it can represent categories that cannot be represented given the resources of S_2, when it can cross-classify the categories represented by S_2, and when it can represent categories that S_2 can, but in an informationally reduced fashion. Abstraction is thus not only a complex property, having several possibly separable components, it is also defined relationally between pairs of systems. Concepts are abstract relative to our perceptual systems, and also with respect to many intermediate nonconceptual processors as well. So the first function of conceptualized thought is to enable the creature that has them to transcend the ways of representing the world that its other cognitive systems give it.

3.2. Causal autonomy

Second, concepts are *causally autonomous*. This means at least two things. First, they are capable of being tokened without the presence of their referents. Crudely, we can think about banana crème pies without being hit in the face with one. This capacity comes in several strengths. In the weak form we can think about the referent where it exists but is simply not perceptually available at the moment. There are no pies here, but I can wish for one and wonder how to make it. In the stronger form, we can entertain concepts of things even despite the fact that their referents do not and perhaps cannot exist. Unicorns, perfect circles, and gods do not exist, hence our concepts of these things are necessarily causally isolated from them. Even so, they have been the subjects of an enormous amount of cogitation. Concepts enable our representational powers to persist across *absences*.

Second, even when the referent of a concept is causally present and impinging on the senses, the way in which we represent and reason about it is in principle independent of our ongoing interactions with it and the world. We may decide to think about the cat that we are looking at, or we may not. And even when we are thinking about it, the way in which we are thinking about it may depart from the way that we perceive it to be (since appearances can be deceiving) or believe it to be (if we are reasoning counterfactually).

What these two characteristics show is that the conceptual system is to some degree *causally autonomous* from how the world affects us perceptually. This idea has been expressed by saying that concepts are under endogenous or organismic control (Prinz, 2002). This autonomy may be automatic, or it may be intentionally engaged and directed. For an example of automatic or unintentional disengagement, consider the phenomenon known as “mind wandering”, in which cognition proceeds according to its own internally driven schedule despite what the creature is perceiving. When there is no particular goal being pursued and attention is not captured by any specific object, trains of thought stop and start based only on our own idiosyncratic associations, interests, dreams, memories, and needs. These also intrude into otherwise organized and focused cognitive processes, depending on one’s ability to inhibit them,. This is an uncontrolled form of causal autonomy.

More sophisticated forms of causal autonomy involve the creature’s being able to direct its own thoughts; for example, to think about quadratic equations rather than pie because that is what the creature *wants* to do. Deciding to think about one thing rather than another is a form of intentional mental action. Hence it is one typical feature of higher cognitive processes that they can be under intentional control or, more broadly, they can be subject to metacognitive governance. Metacognition involves being able to represent and direct the course of our psychological states, a form of causal autonomy.

The capacity to deal with absences through causally autonomous cognition involves “detached” representations (Gärdenfors, 1996; Sterelny, 2003), which can be manipulated independently of the creature’s present environment. Information about particulars needs to be stored and retrieved for future use—I need to be able to locate my car keys in the morning, birds need to be able to locate caches of stored food in winter, and so on. Information about general categories (ways to get a cab in Chicago, kinds of berries that are edible in these woods) also needs to be accessible. Going beyond merely reacting to what the present environment contains requires planning for situations that don’t now exist and may never exist. Planning, in turn, requires the ability to entertain possible scenarios and reason about how to either make them the case or not. And this requires bringing to bear knowledge about the particular and general facts about the world beyond what is immediately present.

This sort of autonomy also shows up in hypothetical reasoning of various kinds. Hypotheses are conjectures concerning what might be the case. While some are straightforward generalizations of experience, others (e.g., abductive inferences) are ampliative leaps that are not determined by the perceptual evidence accumulated thus far. Long-range planning (for where to find food and shelter when winter comes, or what to do in order to get a Ph.D.) demands an especially high degree of causal autonomy. Wherever the goal to be achieved is sufficiently far ahead in time or requires sufficiently many intervening stages, the only way to reliably guide oneself to it is by disengaging from present concerns.

Causally autonomous cognition is closely connected with counterfactual thinking, though they do not precisely come to the same thing. Being able to entertain thoughts that are not causally determined by what is present is part of being able to entertain thoughts of what is not the case. Proper counterfactual cognition requires more than this, of course: the creature needs to be aware (in some sense) that the scenario being entertained *is* one that isn't the case, otherwise there would be nothing stopping it from acting as if it were. Hence the close relationship between the sort of *supposing* that goes on in counterfactual thinking and *imagining*, *pretending*, and *pretense* more generally. These capacities, despite their differences, derive from the same functional template, namely the causal autonomy of cognition, wherein representations are manipulated in a way that is independent of ongoing sensorimotor processing and which is also marked as explicitly nonfactive.

The second function of conceptualized thought, then, is to be available for deployment and manipulation in ways that are driven by the internal condition of the organism and not by the environment or its impact on the organism, where this may be guided by practical ends (as in planning), or by theoretical ends (as in reasoning and generating explanations).

3.3. Free recombination

Third, concepts are capable of free recombination. By this I mean that they obey Evans' Generality Constraint, or something much like it. In Evans' terms, GC amounts to the following: if a creature can think "a is F", and can also entertain thoughts about G, then it can think that "a is G"; and if it can also entertain thoughts about *b*, then it can think that "*b* is F". If you have available an individual concept and a predicate concept, and you are capable of combining them, then you are also capable of exercising that combinatorial ability with respect to all of the other individual and predicate concepts that you have. A child who can think thoughts about the individuals Daddy, Mommy, and Kitty, and who can think "Kitty is mean" and "Mommy is nice" can also think "Daddy is mean" and "Kitty is nice", *inter alia*. Conceptualized thought is general insofar as it allows this sort of free recombination regardless of the content of the concepts involved. Put a slightly different way, once the possible forms of combination are fixed, there are no restrictions on the sort of information that can be accessed and combined by concept possessors.

As stated, GC only applies to "atomic" thoughts, but it is easily generalized. For instance, where P, Q , and R stand for atomic thoughts, if a creature can think "P and Q", then it can also think "P and R". And where it can think "Fs are Gs" and can also think about Hs, then it can think "Fs are Hs". These category-combining thoughts make it clear that a significant function of the Generality Constraint construed as being about *free recombination* is that it allows us to think these sorts of potentially cross-domain thoughts.[7] If we can think that “Sophie is a cat” (a biological classification), and that “cats are prohibited in the building” (an expression of a social norm), then we can think “Sophie is prohibited in the building” (an application of a norm to an individual).

Being able to integrate information across disparate domains in this way is central to the ability to draw productive inferences and reason flexibly about the world, which consists of an overlapping mosaic of properties drawn from indefinitely many interrelated realms. Of course, GC needs to be put together with sufficiently powerful inferential apparatus in order to generate these conclusions. From “Fs are Gs” and “Gs are Hs”, the conclusion that “Fs are Hs” can only be drawn if one is equipped to draw the relevant inference. One needs access to a rule implementing transitive categorical inference. By the same token, causal inferences and other forms of nondeductive reasoning also require premises that bridge representational domains. Having an overall representational system satisfying GC is necessary for this.

Saying that concepts satisfy the Generality Constraint is another way of saying that conceptual representations are to a certain extent context-independent. That is, they are capable of being deployed in connection with any number of different concepts for an open-ended number of purposes, for the satisfaction of indefinitely many tasks and goals, all under the control of a wide array of processes. For example, presumably most creatures are capable of reasoning about mates and possible mating strategies. And at least some employ a theory of mind

[7] I have emphasized the role of GC in allowing cross-domain thought and inference in Weiskopf (2010), where I also argued that massively modular minds cannot conform to GC and hence cannot be concept possessors in this sense. Here one might wonder whether requiring free recombination involves a violation of my neutrality principle; that is, whether this is a way of ruling out massively modular accounts of higher cognition by fiat. I think not, since it needs to be *argued* that massive modularity violates GC; it is not simply built into the view that it does. Free recombination picks out a functional property that is independent of the organization of cognitive architecture and its realization. Whether a particular architecture *can* achieve Generality is an open question, and so no questions are being begged vis-à-vis neutrality.

in achieving their mating ends—they operate with a range of mentalistic construals of the thoughts, desires, and emotions of their conspecifics. And finally, some are capable of thinking about quantity in general, as well as particular numbers such as 2, 3, etc. But only by putting these capacities together could a creature arrive at the baroque combinations of parallel romantic entanglements involved in the average daytime soap opera, in which one needs to manage the (often multiple) romantic affairs of several protagonists engaged in complex webs of deception and manipulation. Putting together the separate domains of mating, minds, and math requires Generality.

The notion of Generality as deployed by Evans ties it closely with the notion of *fully propositional* thought. GC can be seen as a closure condition on the domain of propositions that a system can entertain. But we can envisage forms of the constraint that do not require propositional representation. All that the basic notion of Generality needs is that representations have some form of componential structure. For a non-propositional example, consider a Tinkertoy model of chemical structure, in which spheres represent various types of atoms and sticks represent chemical bonds. The size and shape of a stick determines the type of sphere it can be combined with, and the whole set of possible arrangements can be generated by a set of rules not unlike a grammar. Yet while these chemical models represent possible objects rather than propositions, they conform to Generality in our extended sense: where *a* and *b* are spheres having the same configuration of holes (similar size, shape, etc.), they are intersubstitutable in a model in the same way that individual concepts are intersubstitutable in thoughts. Any process that can manipulate and transform a model containing *a* should be able to do the same for one containing *b*.

A final point: Generality is not to be confused with recursion. Many have pinpointed the emergence of recursion as the decisive transformation that enabled human thought to transcend its evolutionary antecedents.[8] The power of recursion lies in its giving access to an infinite number of new representations from a finite base. All suprasentential inferences, and many subsentential ones, require structures that display at least some recursively embedded structure (e.g., what the propositional connectives deliver). So systems that display free recombination only become truly useful once simple recursion is available. It goes without saying that recursive representational systems obey Generality at least in their recursive parts, since, e.g., any propositional representations can be intersubstituted for *p* and *q* in the material conditional *if p then q*. This does not strictly require imposing Generality on the atomic propositions themselves, although dissociating these two properties would be strange. For present purposes, we can regard recursive systems as a special case of Generality. Put differently, free recombination is a more cognitively fundamental property than productivity.

So the third function of conceptualized thought is to enable the free and context-independent combination of information of information across domains, and more broadly to facilitate the construction of complex representations that can be the subjects of a range of inferential processes.

4. Higher cognition as a psychological kind

These three properties constitute the core functions of the human conceptual system. These properties are graded, so there will be cases of marginal concept possession involving

[8] See Corballis (2011) for a defense of recursion as central to explaining human cognitive uniqueness. There have also been questions about just *how* recursive human thought is—particularly as expressed in language (Evans & Levinson, 2009; Everett, 2005).

creatures that have these qualities only to a certain degree; but more importantly, these properties pick out functional roles that are logically independent and empirically dissociable as well.

Conceptualized thought, perhaps uniquely, occupies the apex defined by joint possession of all three properties, with higher faculties as conceived of here occupying a graded series, on which some faculties may be higher in one respect without being so in another. This fractional approach to describing higher faculties also avoids the problems posed by earlier single-criterion accounts.[9] It overcomes the Unity Problem by showing that higher faculties share core similarities in their causal role. It overcomes the Richness Problem by providing a set of possible distinctions among higher faculties, each of which is itself graded in various ways, thus generating a potentially elaborate taxonomy. It avoids the Collapse Problem by adding substantially new functionality to the cognitive system at each level of the hierarchy. Finally, it sustains Neutrality by maintaining silence about the underlying systems-level decomposition into modular vs. nonmodular components, as well as on the precise neural substrates and evolutionary origins of higher faculties.

On this picture we can see both the hierarchy of higher cognitive faculties generally, and conceptualized thought in particular, as constituting psychological kinds. A *psychological kind*, in the present sense, is a type of structure that plays an important unifying or explanatory role in

[9] Several existing accounts of higher cognition are similar to the present one, which attempts to build on their insights. Allen & Hauser (1991) argued that higher cognition in animals involves being able to represent abstract categories, that is, those that have highly various perceptual manifestations. This is close to the representational abstractness criterion proposed here. Christensen & Hooker (2000) give a highly detailed account of how various forms of causal autonomy might be structured and related to lower-level capacities for maintaining the integrity of living forms. Camp (2009) argues that both Generality and stimulus-independence are required for conceptual thought, but does not separately mention representational abstractness. And Sterelny (2003) proposes that robust tracking using decoupled representations is essential for belief-like cognition; this corresponds to the combination of causal autonomy and representational abstractness. Another extraordinarily detailed proposal concerning how these two might be fused comes in Amati & Shallice's (2007) theory of abstract projectuality. On the present view, all of these make a distinctive and necessary contribution to higher cognition and conceptual thought proper, though none are the whole story taken on their own.

a sufficiently wide range of cognitive and behavioral phenomena (Weiskopf, 2011). The taxonomy itself meets this standard insofar as it lays out three well-defined and scientifically interesting dimensions of variance, and provides a framework for guiding investigations into cognitive development and evolution, as well as in comparative cognitive ethology.

First, these capacities play a role in developmental explanations. Developmental psychologists are keenly interested in characterizing the nature and timing of the so-called "perceptual-to-conceptual shift" (Rakison, 2005). This refers to the period when infants go from being able to perceptually track complex categories to having genuine concepts of them. Even very young infants can make perceptual discriminations among relatively high-level categories, such as animals, vehicles, or furniture. But they are initially able to do so only in passive tasks such as perceptual habituation. It is widely agreed that this evidence by itself is too weak to establish that they have the relevant concepts. But there is disagreement over what more is needed.

On this account, two things are required: the ability to carry out at least somewhat sustained off-line processing using these categories, and the ability to combine these categories with others. Without some such account of the functional role of higher thought, we cannot frame experiments or test hypotheses aimed at discovering when such thought comes on-line. With such an account, however, we can do so—and, importantly, in a way that delivers knowledge about infants' capacities but does not require detailed knowledge about the *precise* form of the underlying systems and representations they employ. Relatively high level functional categories often have this feature: we have some idea of what is involved in their presence even if we do not know just how, in detail, they are implemented.

Second, these capacities may have distinctive evolutionary origins. In one of the most careful and sustained discussions of this question, Sterelny (2003) proposes several different types of environmental structure that may impose selection pressures towards the development of abstract representations that can be decoupled from ongoing input. Importantly, these pressures have somewhat different sources. The former depend on whether the categories necessary for a creature's survival are "transparent" to its sensory systems (pp. 20-26), while the latter depend on the kind of cognitive tracking that emerges mainly among social animals with an interest in reading and manipulating each other's thoughts (p. 76). It remains to be investigated what sorts of pressures, if any, facilitate the development of freely combinatorial representational vehicles. Generally, however, the fact that a capacity is potentially subject to such selection pressures indicates not only that it can be explained, but also that its presence can in turn explain the success facts about an organism in its environment.

Third, note that these capacities are *polymorphic*. By this I mean that they are general types of cognitive capacities—functional templates—which may be instantiated in cognitive systems in many different ways. There are many different low-level organizations of cognitive components that can satisfy the requirements for various forms of higher cognition. These components may differ in their representational resources, processing characteristics, overall location in the architecture, and so on, so that creatures that have similarly sophisticated capacities may still differ from one another in their precise internal organization (their cognitive "wiring diagram").

However, we may nevertheless want to state generalizations across creatures having different forms of these polymorphic capacities.[10] To take an ethological example, there may be

[10] This dovetails with the earlier emphasis on neutrality: in merely stating the functional requirements on higher cognition, we do not want to *presuppose* the existence of one or another type of underlying neural or systems-level

little reason to suppose that higher cognition in nonhuman animals precisely resembles that in humans. Different sensory capacities, needs, and environments may interact to lead them to make radically different conceptual divisions in their worlds, which in turn leads to higher capacities that deal in distinctive types of rules and representations. Indeed, for all we know there may be substantial diversity in these capacities even among humans considered developmentally and cross-culturally.

Given this diversity, we may still want to co-classify these beings that possess differently structured forms of these polymorphic capacities. While they may not reason about or represent the world in precisely the same way that we do, they are undeniably doing something of the same broad type. In order to isolate this capacity, trace its origin, and compare its structure and content to our own, we need a way of describing it that is a level of abstraction removed from the precise details of how it processes information. I would emphasize that there is nothing unusual in this: precisely the same questions would arise if we were investigating alien neurobiology and attempting to determine what sorts of neurons, neurotransmitters, synaptic junctions, etc., they possessed. Without an appropriate functional taxonomy in hand, such comparative investigations have no plausible starting point.

What it means for higher cognition to be a kind is just for the division between higher and lower systems to capture categories that are structurally real and explanatorily fruitful. Taxonomic schemes generally pay their way in science by unifying disparate phenomena, by helping to frame new targets for explanation, and by suggesting novel empirical hypotheses. This

---

organization. That does not mean, however, that certain kinds of internal organization won't be ruled out by these requirements. To take an example, subsumption-style architectures such as those used in certain types of mobile robots may be incapable of satisfying the requirements on higher cognition, even though they possess robust behavioral capacities. Since this architectural restriction is discovered rather than stipulated, it does not violate neutrality.

brief sketch should suggest that the hierarchy of thought as conceived of here can do all three of these things.

5. Conclusions

One might, not unreasonably, have supposed that the very idea of dividing cognitive capacities into higher and lower reflects a kind of philosophical atavism. Neither the casual way in which the distinction is made by psychologists, nor the poverty of standard attempts to unpack it inspire hope. However, it turns out that not only can we make a rather finely structured set of distinctions that captures many of the relevant phenomena, these distinctions also seem to correspond to genuinely explanatory psychological categories. Seeing organisms through this functional lens, then, allows us to gain empirical purchase on significant developmental, evolutionary, and ethological questions concerning their cognitive structure.

Acknowledgements

Thanks to those who commented on an earlier version of this paper during the online conference: Mark Sprevak, Bryce Huebner, Justin Fisher, Robert O’Shaughnessey, and Georg Theiner. Their comments and discussion helped me to think through these issues and clarify the paper at several points. Thanks also to the editors, Mark Sprevak and Jesper Kallestrup, for their kind invitation to participate in this volume.

References

Allen, C., & Hauser, M. (1991). Concept attribution in nonhuman animals: Theoretical and methodological problems in ascribing complex mental processes. *Philosophy of Science, 58,* 221-40.

Allen, J. S. (2009). *The Lives of the Brain*. Cambridge: Harvard University Press.

Amati, D., & Shallice, T. (2007). On the emergence of modern humans. *Cognition, 103,* 358-85.

Anderson, M. L. (2010). Neural reuse: A fundamental organizational principle of the brain. *Behavioral and Brain Sciences, 33,* 245-313.

Barrett, H. C. (forthcoming). A hierarchical model of the evolution of human brain specializations. *Proceedings of the National Academy of Sciences*.

Beck, J. (forthcoming). The generality constraint and the structure of thought. *Mind*.

Benton, A. L. (1991). The prefrontal region: Its early history. In H. Levin, H. Eisenberg, and A. Benton (Eds.), *Frontal Lobe Function and Dysfunction* (pp. 3-12). Oxford: Oxford University Press.

Camp, E. (2009). Putting thoughts to work: Concepts, systematicity, and stimulus-independence. *Philosophy and Phenomenological Research, 78,* 275-311.

Carey, S. (2009). *The Origin of Concepts*. Oxford: Oxford University Press.

Carruthers, P. (2004). On being simple-minded. *American Philosophical Quarterly, 41,* 205–22.

Carruthers, P. (2006). *The Architecture of the Mind*. Oxford: Oxford University Press.

Christensen, W. D., & Hooker, C. A. (2000). An interactivist-constructivist approach to intelligence: Self-directed anticipative learning. *Philosophical Psychology, 13,* 7-45.

Coltheart, M. (1999). Modularity and cognition. *Trends in Cognitive Science, 3,* 115-20.

Corballis, M. C. (2011). *The Recursive Mind: The Origins of Human Language, Thought, and Civilization*. Princeton: Princeton University Press.

Danziger, K. (1997). *Naming the Mind: How Psychology Found its Language*. London: Sage Publications.

Elman, J. L. (1993). Learning and development in neural networks: The importance of starting small. *Cognition, 48,* 71–99.

Evans, N., & Levinson, S. (2009). The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and Brain Sciences, 32,* 429-92.

Everett, D. (2005). Cultural constraints on grammar and cognition in Pirahã. *Current Anthropology, 46,* 621-46.

Finger, S. (1994). *Origins of Neuroscience: A History of Explorations into Brain Function*. Oxford: Oxford University Press.

Gärdenfors, P. (1996). Cued and detached representations in animal cognition. *Behavioral Processes, 35,* 263-73.

Ghazanfar, A. A., & Schroeder, C. E. (2006). Is neocortex essentially multisensory? *Trends in Cognitive Science, 10,* 278-85.

Hatfield, G. (1997). Wundt and psychology as science: Disciplinary transformations. *Perspectives on Science, 5,* 349-82.

Knill, D. C., & Richards, W. (Eds.) (1996). *Perception as Bayesian Inference*. Cambridge: Cambridge University Press.

Kveraga, K., Ghurman, A. S., & Bar, M. (2007). Top-down predictions in the cognitive brain. *Brain and Cognition, 65,* 145-68.

Neisser, U. (1967). *Cognitive Psychology*. New York: Meredith.

Preuss, T. M. (2011). The human brain: Rewired and running hot. *Annals of the New York Academy of Science, 1225,* 182-91.

Prinz, J. (2002). *Furnishing the Mind*. Cambridge: MIT Press.

Rakison, D. H. (2005). The perceptual to conceptual shift in infancy and early childhood: A surface or deep distinction? In L. Gershkoff-Stowe & D. H. Rakison (Eds)., *Building Object Categories in Developmental Time* (pp. 131-58). Mahwah, NJ: Lawrence Erlbaum.

Rock, I. (1983). *The Logic of Perception*. Cambridge: MIT Press.

Schroeder, C. E., & Foxe, J. J. (2004). Multisensory convergence in early cortical processing. In G. A. Calvert, C. Spence, & B. E. Stein (Eds.), *The Handbook of Multisensory Processes* (pp. 295-310). Cambridge: MIT Press.

Sherwood, C. C., Subiaul, F., & Zawidzki, T. W. (2008). A natural history of the human mind: Tracing evolutionary changes in brain and cognition. *Journal of Anatomy, 212,* 426-54.

Sterelny, K. (2003). *Thought in a Hostile World*. Malden, MA: Blackwell.

Stich, S. (1978). Beliefs and subdoxastic states. *Philosophy of Science, 45,* 499-18.

van Gelder, T. (1993). Is cognition categorization? *Psychology of Learning and Motivation, 29,* 469-94.

Weiskopf, D. A. (2010). Concepts and the modularity of thought. *Dialectica, 64,* 107-30.

Weiskopf, D. A. (2011). The functional unity of special science kinds. *British Journal for the Philosophy of Science, 62,* 233-58.