Are there psychological species?

Joshua Fost Ph.D.
Philosophy Dept., Neuberger Hall
Portland State University
Portland, OR 97207-0751
jwfost@pdx.edu

Fost JW (2014) Are there psychological species? *The Review of Philosophy and Psychology.* doi: 10.1007/s13164-014-0227-y.

The final publication is available at Springer via http://dx.doi.org/10.1007/s13164-014-0227-y

---

Abstract
A common reaction to functional diversity is to group entities into clusters that are functionally similar. I argue here that people are diverse with respect to reasoning-related processes, and that these processes satisfy the basic requirements for evolving entities: they are heritable, mutable, and subject to selective pressures. I propose a metric to quantify functional difference and show how this can be used to place psychological processes into a structure akin to a phylogenetic or evolutionary tree. Three species concepts are repurposed from biology and used to understand relationships in that tree.

Key words: psychological types; epistemic rationality; functionalism; species concepts; philosophy of psychology

1. Introduction

In January 1962, Bertrand Russell sent the following letter to Sir Oswald Mosley, founder of the British Union of Fascists:

> Dear Sir Oswald,
> Thank you for your letter and for your enclosures. I have given some thought to our recent correspondence. It is always difficult to decide on how to respond to people whose ethos is so alien and, in fact, repellant to one's own. It is not that I take exception to the general points made by you but that every ounce of my energy has been devoted to an active opposition to cruel bigotry, compulsive violence, and the sadistic persecution which has characterised the philosophy and practice of fascism.
>
> I feel obliged to say that the emotional universes we inhabit are so distinct, and in deepest ways opposed, that nothing fruitful or sincere could ever emerge from association between us.
>
> I should like you to understand the intensity of this conviction on my part. It is not out of any attempt to be rude that I say this but because of all that I value in human experience and human achievement.
>
> Yours sincerely, Bertrand Russell.

Lord Russell gives voice to an unsettling but probably familiar sensation. Unbridgeable chasms of interpersonal difference appear perhaps most commonly in religion and politics, but also in ethics, aesthetics, even humor. One finds sublime what another finds crass; one finds impossible what another finds self-evident; one weeps and another guffaws. In such cases it will not do to say merely that the characters involved have varied opinions or tastes or that they reasoned similarly but reached different conclusions. Rather, we must say—indeed we do say—that the manners in which the two people think and feel and judge are so incompatible that they may as well live on different planets. They may as well belong to different psychological species.

My purpose in this paper is to make a case for the utility of translating species concepts from biology and applying them to psychology. More specifically, I will suggest that the processes associated with epistemic reasoning, i.e. those determining what we believe to be true, can be and ought to be separated into groups according to the some of the same classificatory tools used to draw species boundaries in biology. I will also suggest that interpersonal variability in these processes constitutes a better basis for typological distinctions than many of the psychological properties presently used for that purpose, such as personality traits. My argument, in brief, is as follows (of course some of the terms here will require and receive explication later):

P1: Reasoning-related processes are structurally diverse.
P2: Minds implement reasoning-related processes. Ergo, people are diverse with respect to reasoning-related processes.
P3: Reasoning-related processes are heritable, mutable, and subject to selective pressures. Ergo, they evolve.
P4: Species concepts clarify relationships amongst evolving entities.
C: Species concepts clarify the diversity seen in reasoning-related processes.

The paper will develop the argument in that sequence. P1 is implicitly a definition of "structural diversity," so that will come first. P2 is a special case of the computational theory of mind, which I will assume but not defend. Defense of P3 and P4 will constitute the bulk of the paper. Upon concluding my argument, I will review several other classificatory systems in psychology and consider whether they warrant a similar application of species concepts. Finally, I'll close with a few thoughts on possible causes of psychological speciation.

2. Structural and Functional Diversity

My first task is to clarify what I mean by "structural diversity" in the domain of reasoning-related processes. Making a structural discrimination is a very ordinary, not to say unsubtle, conceptual move. Toasters and automobiles are obviously different things. They have different parts and do different things to bread and gasoline. Frogs and antelopes are also obviously different structurally and functionally: understandably so because evolution, like human invention, explores structural and functional design space.

To say that people differ in the relevant structural and functional ways is to assert something about what parts they have and how the parts interact. At the behavioral level,

people are obviously diverse on reasoning-related tasks: they give different answers to the same question. By itself, however, this doesn't settle whether their reasoning engines differ: perhaps they understood the task differently, were differently affected by circumstances of the testing environment, or were differently motivated to perform well. This is the *competence vs. performance* debate. If interpersonal differences were largely performance-based, there would be little point in segregating people into categories, species-like or otherwise, because people might be exposed as rationally identical if we only controlled for all non-rational factors. My arguments below will only be relevant if a sizeable fraction of rational diversity turns out to stem from competence differences, so let me present some reasons for thinking that this is the case. First, however, a terminological caveat: the word "competence" suggests a particular view of rationality, namely a *normativist* view in which there is such a thing as a "competent" reasoner, and anyone who reasons differently is to some degree incompetent. With respect to epistemic rationality, I consider myself a normativist, but for my purposes in this paper it won't matter. I will be arguing that there are differences between people and that we can make sense of these differences with species concepts. I needn't enthrone any particular pattern of reasoning as being normative for this project—all I need is for competence differences to exist.

Perhaps the most straightforward evidence for competence difference is the fact that some people, such as logicians and statisticians, consciously use formal symbolic processes to determine some of their beliefs. Other people are not even aware of these procedures (nor do they apply them unconsciously) and as a result they arrive at different beliefs. Awareness (or neglect) of base rate is an example. Consider the now classic illustration: a medical test is 99% accurate (that is, 99% of the people with the disease test positive, and 99% of healthy people test negative) for a disease that is rare: 1% of the population actually has it. Suppose a patient tests positive. What is the probability that s/he actually has the disease? Many people answer 99% or thereabouts. The correct answer is 50%. If one knows and applies Bayes' rule, this is easy. If one does not know Bayes' rule, coming up with the correct answer is much more difficult, and many people will insist that the 50% number cannot possibly be right. More generally, the heuristics and biases literature contains "dozens of empirical studies [that] have shown that there are few tasks…where all untutored laypersons give the same response" (Stanovich, 2012) but where some (even untutored) subjects give the response traditionally construed as normative. Between the normative responders and the untutored laypersons is a difference that reappears predictably on problems of the same type. That predictability is hard to attribute to a performance difference, because the circumstances affecting performance presumably vary from experiment to experiment.

Consistent with this view of competence differences, some authors write in terms of intuitive vs. reflective thinking *dispositions* (Svedholm & Lindeman, 2013), decision-making *styles* (Dewberry, Juanchich, & Narendran, 2013) and "deliberative [vs.] intuitive *thinkers*" (my emphasis) (Mata, Ferreira, & Sherman, 2013). Other authors take a developmental approach, investigating when various reasoning competencies first appear (Santamaría, Tse, Moreno-Ríos, & García-Madruga, 2013). This automatically implies that some people (i.e. young children) lack certain rational competencies altogether, so between them and more mature subjects there is a competence difference. Yet other investigators are studying the brain areas responsible for instantiating some of these

reasoning processes (Corradini & Antonietti, 2013; Essex, Clinton, Wonderley, & Zald, 2012; Reverberi et al., 2012), so again, different brain structures might produce different competencies.

What *are* these reasoning-related processes, specifically? We see them indirectly. Heuristics like Kahneman's "what you see is all there is" and biases like base-rate neglect, confirmation bias, and the gambler's fallacy suggest that there is a process, or a set of processes, concerned with the collection of relevant data. Social assessments such as the appraisal of a source's credibility or the popularity of a belief also figure in to data selection. A different set of heuristics might include things like conservatism ("if I already believe it it's probably true"), the assumption that "cognitive ease equals truth," and "if I've heard it before it's more likely to be true." These examples suggest a set of processes concerned with determining the degree to which propositions harmonize with existing beliefs and gilding those that do. These are all fast and largely automatic (System 1) examples, but more deliberative analytical (System 2) processes may also play a part in the reasoning process: deductive and inductive logic, probability and statistics, and so on.

I do not claim that every one of these processes is functionally separable from the others, or from more general psychological processes, like attention or memory. It's possible that some are separable in that way, and proponents of the massive modularity hypothesis would likely agree. Peter Godfrey-Smith (2002) agrees that "cognitive machinery is very diverse," that this machinery should be thought of as a collection of capacities that are realized in different ways, and includes under this umbrella many non-human and even non-animal forms of adaptive responding. While I am amenable to extending the present arguments to this much broader scope, doing so here would take us too far afield. So at the moment, I only want to claim that (a) epistemic rationality appears to have parts—it is not a unitary, holistic capacity, and (b) not everyone has the same parts…or at least, not everyone has their parts working together in the same way.

So there are good reasons to suppose that people are diverse in their rational competencies. What do I mean by saying that these correspond to *structural* differences? Briefly, what I mean is that there is a physical basis for the diversity. The structures I have in mind are neural (Godfrey-Smith thinks of it more broadly), but I want to illustrate the point first with a non-neural example, in part because the quantitative measure of structural difference that I'll be introducing later is clearer here.

The example I have in mind is sickle cell disease. The disease is caused by a point mutation of the beta-globin gene *HBB*. The relevant function of this gene is the oxygen carrying capacity it confers upon red blood cells, and it is impairment of this function that constitutes the disease. Phenylketonuria, cystic fibrosis, and many other point-mutation genetic disorders have similar stories: a simple physical change in a gene degrades the function. In sickle cell disease there are other factors that contribute to oxygen carrying capacity, but in severe phenylketonuria the impairment caused by the mutation is sudden and total: the mutant enzyme does not catalyze the target reaction at all and there is no other catalyst to do the job.

In sickle cell and phenylketonuria, then, we have three features to attend to: (a) a function of interest (the carrying of oxygen, the breakdown of phenylalanine), (b) a physical entity that implements that function (hemoglobin, phenylalanine hydroxylase),

and (c) a specification for the creation of that entity (DNA). Changes in the specification cause differences in physical structure and, thereby, the function.

With respect to *psychological* functions or processes, the situation is similar. Each function is implemented by a neural circuit, so it is there (rather than in protein structure) that we can cash out the physical attributes in virtue of which two processes are distinct. A circuit with property P implements function F, but circuits without this property do not. (I will use the famous example of the artificial neural network known as the *perceptron* to illustrate this more fully below.) So in both the genetic and psychological cases, there is a function or process of interest, and a physical structure that implements that function. In the last feature, however, the *specification* for the physical structure, there are some differences between psychology and genetics. Whereas DNA contains the complete recipe for a protein, it (DNA) does not contain the complete recipe for the neural circuit—at least not for most of the complex cognitive processes under consideration here. That is, most neural circuits are probably shaped much more by experience than by a genetic encoding. The existence of a specification, though, is important for my argument because, as I'll explain below, it will provide the means of quantifying the distance between two functions. Where, then, shall we turn for a specification of the neural structure?

If the computational theory of mind is right (and via P2 I am assuming that it is) then *every psychological function is equivalent to a Turing machine and therefore encodable in a static information-bearing structure* (Turing, 1937). In practice, such structures are just strings of symbols that, when interpreted by a Universal Turing Machine, instantiate a particular Turing machine and thereby a particular computational process. One can imagine such a Turing machine simulating the processes of the brain/mind at whatever level is needed, from the extremely tiny molecular dynamical level up through the symbolic, representational level. It doesn't matter that such encodings don't appear anywhere in the human organism in the way that DNA does for genes and protein-based functions. I will make use of the encodings as a computational convenience: it is hard to see how to rigorously characterize the difference between two rational processes, but it is rather easy to see how to compare two strings. Since there is a one-to-one correspondence between the strings and the processes, we can use the strings to precisely characterize—quantify, actually—the difference between processes. Here is how it works.

A metric known as the *edit distance* (Levenshtein, 1966) counts how many symbols must be inserted, deleted, or substituted to transform one string into another. So "house" and "horse" are separated by an edit distance of one, while "chair" and "charm" are separated by an edit distance of two (delete the 'i', then add the 'm'). This and similar metrics are commonly used to infer phylogenetic relationships from DNA sequences (Nei, 1987) because DNA sequences are, like Turing machine encodings, just strings of symbols. Using the edit distance metric, the space of all Turing machines can be meaningfully organized as a genealogical tree (Fig. 1). Starting with one arbitrary reference encoding at the root, we add as children all the encodings that are exactly one edit away. We repeat this procedure for each of those children, until the tree contains every possible computable process up to some encoding length limit. In the context of this paper, think of each node as a set of epistemic reasoning processes potentially realizable by a person's brain.

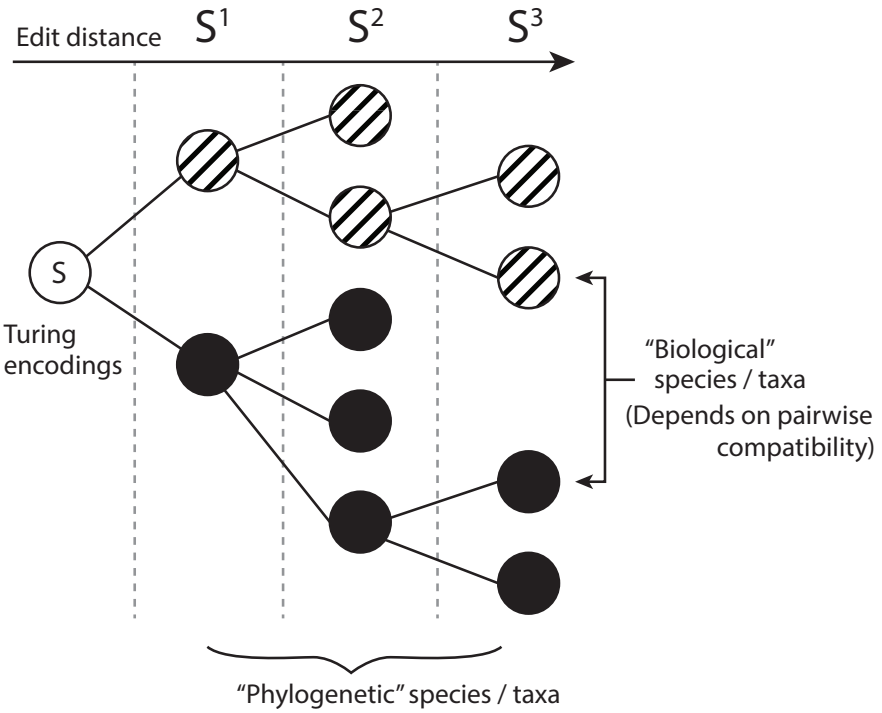

Figure 1. A genealogical tree of Turing encodings based on the edit distance. S^n is the set of encodings edit distance *n* away from the reference encoding *S*. Note that different encodings (shown here as hatched or solid based on differing ancestry) might be weakly equivalent in Pylyshyn's sense, i.e. compute the same answer in different ways. When this is true, they generate no internal contradiction, so would count as conspecific under the repurposed biological and ecological species concepts but not the phylogenetic concept.

Let's take stock of the four entities under discussion: (a) a psychological process, i.e. a series of operations on mental representations; (b) a Turing encoding for that process; (c) a neural circuit that implements that Turing machine; and (d) a person whose brain contains that circuit. With respect to the processes themselves, we can talk of "functional difference"—which is what's on display when two people give different answers to the same question—and with respect to the neural or algorithmic underpinnings of those functions, we can talk of "structural difference"—which is what's on display when two people have different brain structures or Turing encodings for their reasoning process, irrespective of behavior. This difference reflects Pylyshyn's (1984) distinction between weak and strong functional equivalence. Strong equivalence is the same function implemented by the same algorithm (in this case, the same Turing machine), while weak equivalence is the same function implemented by different algorithms. When I talk about reasoning-related differences between people, I am concerned with structural differences that produce functional differences.

Let me pause here to address a possible concern, which is that there is so large a gap between Turing encodings of microphysical processes, on the one hand, and epistemic reasoning on the other, that the scheme I am constructing is just too opaque to assess. How could a Turing encoding of biophysics capture whether someone is subject to the base rate fallacy on a particular problem?

My reflexive reply is that this concern is nothing more than an echo of intuitions against the computational theory of mind. But I am aware that this reply will not appease anti-functionalists, so in the interest of establishing *some* systematized means of measuring interpersonal differences, I will offer another more sympathetic and probably a more productive answer. This one trades on the possibility of a middle ground, a functional level between neurons and behavior that allows us to see more clearly how two systems could be functionally distinct. What I have in mind here are cases in which the macroscopic functioning of system obeys a syntax.

An obvious example is language: parts of the brain implement a grammar that screens for particular input or output patterns. If I hear "John gave Mary the _____," my brain implements a function signaling when the blank is filled in properly (i.e. with a noun phrase) or not. How it does so is not important: what matters is that there is a psychological process that enforces a standard. If process *A* permits an input or output that *B* forbids, we can say that *A* and *B* are functionally distinct. Linguists use such similarities and differences between grammars to infer relationships between languages, constructing the same sorts of evolutionary trees that I support for reasoning.

There are non-linguistic syntactic systems as well: Buzsáki (2010) describes how "neural syntaxes" recombine a finite set of elemental symbols into words, phrases, and sentences. He describes how a number of systems implement such combinatorial syntaxes: odorant coding in the insect brain, birdsong [see also Hanuschkin, Diesmann, & Morrison (2011)], and fixed action patterns like grooming. In each of these cases, mental representations are instantiated by patterns of activity distributed across neural ensembles and syntax is enforced via the combination of synaptic weights and the temporal dynamics of individual neurons that make them responsive to some incoming patterns of activity but not others. If the weights and dynamics are such that a given neuron does not respond to an incoming pattern, this constitutes a disallowed state transition and thereby a syntactic violation. If one system permits a pattern that another forbids (think of two bird species with different templates for mating calls) then they qualify as functionally distinct.

I'll give one last illustration of a functional difference that can be characterized without Turing encodings. Consider the classic example of the single- and multi-layer perceptron, which I mentioned earlier (see Figure 2).

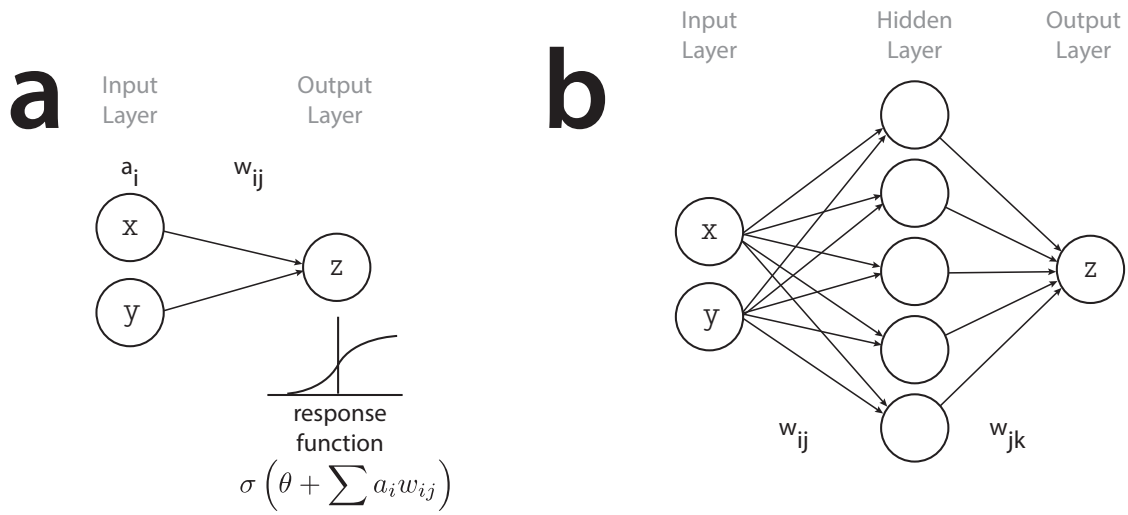

Figure 2. The single- and multi-layer perceptron. (a) The single-layer version cannot compute $z = XOR(x,y)$ regardless of parameter values. It is therefore functionally distinct from (b) the multi-layer version and would qualify as being a different psychological species.

The perceptron was introduced by Rosenblatt (1958) and became an exemplar for early computational neuroscience. The model featured many salient properties of neural systems, such as rich connectivity between neurons, variable synaptic strengths, thresholds, and nonlinear response functions. Rosenblatt showed that even simple devices of this type—single-layer perceptrons—could be trained, by changing the weights of the synapses, to classify complex multi-dimensional stimuli. As described by Minsky & Papert (1969), however, the single-layer machine has shortcomings. It is incapable, for instance, of implementing the exclusive or (XOR) function. That is, there is no set of parameter values (synaptic weights, thresholds, response function, etc.) that will allow it to determine when one, but not both, of two inputs are active. By adding an additional layer, however, thereby creating a *multi-layer* perceptron [see Figure 2(b)], it becomes easy to find a set of values that implement XOR. This discrepancy between the single- and multi-layer perceptron is the sort of structural and functional difference that I have in mind for reasoning.

Even without Minsky & Papert's demonstration, casual inspection of the architectures of the two networks reveals an obvious difference. What is *not* obvious is whether, when, or why that difference matters. There are innumerable variations of neural network architectures, and in general a great deal of work is necessary to ascertain whether two arbitrarily chosen architectures are meaningfully different at the functional level. The force of Minsky & Papert's proof was its rigorous demonstration that the difference between the single- and multi-layer perceptrons *did* matter at the functional level. No matter how a single-layer perceptron is configured, it can never compute XOR.

Taken all together, these examples illustrate what I mean by functional and structural incompatibility. Distinct neural architectures can implement incompatible functions. At the behavioral level, different people reason differently and arrive at different beliefs. I conclude that some of those reasoning differences reflect functional incompatibilities in their reasoning-related processes. If we knew more about how their

brains worked, we could go further to see those differences reflected at the neural and sub-neuronal level, ultimately quantifying those differences using the edit distance on Turing encodings.

That ends my explication of functional and structural distinctions in reasoning-related processes. The next step is to justify an evolutionary approach to understanding diversity in this domain by showing that rational processes satisfy the necessary conditions for evolution, namely that they are mutable, heritable, and subject to selective pressures.

3. The Evolvability of Rational Processes

The entities populating the tree of Figure 1 are unchanging strings, so they fail to satisfy the first condition: mutability. But I am not claiming that Turing encodings *qua* representations of timeless abstract computations are the things that change. What changes, rather, is the particular reasoning process instantiated in a person's brain. The computational theory of mind sanctions an identity relation between a brain and one of the Turing encodings shown in the tree, so as a person's brain changes as a result of biological maturation and learning, they are traversing the tree in some way.

This is not to say that all trajectories through the tree are possible, or equally likely. Brain development is regulated in specific ways: cells multiply in different places at different times, trophic factors regulate macroscopic area-to-area connectivity, synaptic pruning eliminates weak connections, and so on. The frontal cortex, strongly implicated in reasoning, is still maturing well into adolescence. These developmental programs unfold in patterns typical to *Homo sapiens*, so most normal brains proceed along paths which, though not identical, certainly bear a family resemblance. Of course the reason for this commonality is that the genes that guide the programs conferred an adaptive value on our ancestors: they set the stage, as much as genes can, in preparation for the finer-grained learning that was to follow. Presumably much the same thing happens with the developmental programs that build, among other things, the motor control circuits that prepare us for bipedal locomotion but which have to adjust for variable morphology, and the language circuits which prepare us to learn *some* language but do not, indeed cannot, encode a *particular* language.

Let me draw this out with a reasoning-specific example. The expectation of object permanence appears naturally around age 7 months and appears to be the result of functional maturation of frontal lobe circuits (Baird, 2002). Functional maturation includes genetically-driven anatomical development and sensory experience-driven neural fine-tuning. The assumption that solid objects do not simply vanish from the universe constitutes an axiom of epistemic reasoning that adults use all the time, and any individual with a certain set of alleles will make it. Thus, the axiom of object permanence is probably a genetically heritable component of the reasoning process. In contrast, no untutored person, regardless of genetic inheritance, will spontaneously calculate t-statistics to determine whether two sample datasets were drawn from the same population. That reasoning process is not genetically heritable, nor will it emerge *ex nihilo* through typical sensory experiences. It can only be acquired through cultural transmission.

So at least some of an individual's complement of rational processes is heritable (see Fig. 3). Our native brain microstructure, which constitutes our starting point in the

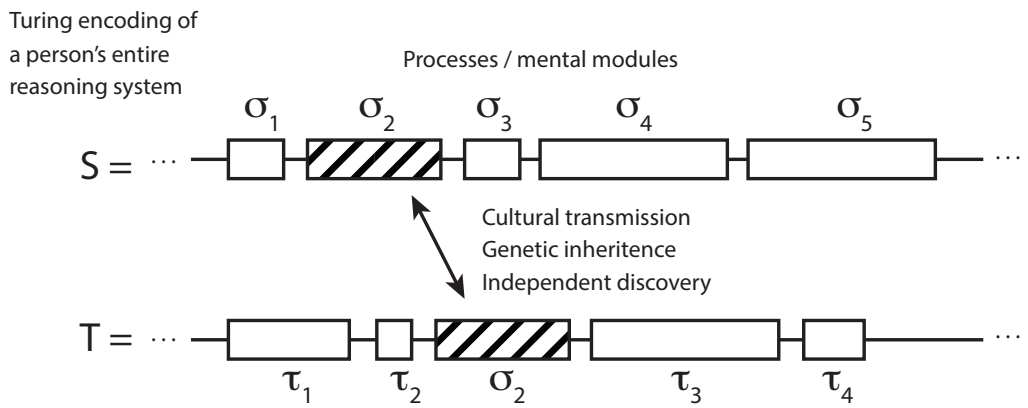

Figure 3. Heritability of reasoning processes. A person's entire complement of reasoning processes is a Turing encoding, e.g. *S* or *T*. This encoding is a string composed of a concatenation of substrings $\{\sigma_i\}$ each of which corresponds to a part of the reasoning process. Some of these processes will be shared between people, and some are genetically or culturally heritable. During learning, the neural networks realizing the σ_i change—adding neurons to or removing them from a network, modifying synapses, or reconfiguring intrinsic cellular parameters. Thus the computation instantiated by the individual changes and they move through the tree of encodings, which may entail a change in species membership.

tree of Turing encodings, is at least in part *genetically* heritable (with the remainder due to developmental influences). Heritability through *cultural* transmission is also possible and is in some ways more powerful than genetic: it is more fine-grained and more flexible in that it allows for the horizontal (within a generation) and vertical transfer of reasoning processes.

To underwrite the conception of rational processes as evolving entities, the last piece required is a demonstration that these processes compete on a fitness landscape, with a selection process that results in increased prevalence for the fitter variants. For a gene, fitness equals fecundity, but self-replication is not the only way for an entity to become more prevalent. Viruses do not self-replicate (nor, ostensibly, do memes…more on this below) but they can still increase in numbers if they do their job well...the job being, in their case, to commandeer the reproductive machinery of a more complex being. So what is the job of a epistemic reasoning process? –or, more to the point, under what circumstances will such a process have a higher likelihood of being transmitted to another person?

I said that there were two ways of inheriting rational processes: genetic and cultural. These two channels, respectively, admit two kinds of circumstances for improved transmission: (a) when the process increases a person's Darwinian fitness, and (b) when the process increases a person's subjective impression of sense-making. The first of these ought to be uncontroversial because it is conventional natural selection: if a gene or set of genes produces a phenotype that increases an organism's fecundity, the genes responsible increase in frequency. The example of object permanence works well here: some genes encode this piece of the reasoning process, and by doing so they increase the odds of organismal survival. Thus the axiom of object permanence has spread through human populations.

The second part is subtler. Why would the mere *impression* of sense-making, rather than *actual* sense-making (whatever that may be), increase the probability of a process's transmission to another person? Why, in other words, doesn't it matter if a rational process conforms to epistemically normative knowledge acquisition? Well, because all that is required for a person's endorsement, whence cultural transmission of a rational process, is their subjective satisfaction with it. Whether the process in question maximizes (for example) the ratio of justified true beliefs to false ones is only relevant if that maximization feels good enough to make a person pass it on.

Consider the epistemic and justificatory structure of conspiracy theories. Canonical cases go like this: I believe *x*, and I also believe that a shadowy group has conspired to suppress evidence for *x* and mislead everyone about it. Now if I find evidence consistent with or uniquely supportive of *x*, my belief in *x* is strengthened. If I fail to find evidence for *x*, or encounter evidence against it, my belief in the conspiracy is strengthened and belief in *x* is preserved. Failure to find evidence for the conspiracy is interpreted as positive evidence for their diabolical efficacy in covering their own tracks, and finally, owing to the epistemic thorniness of negative existential claims, there is no practical way to find positive evidence that the conspiracy doesn't exist.

From this almost impenetrable structure, conspiracy theories help their believers make sense of almost any new data in theory's scope. If I am trying to settle my beliefs about the JFK assassination, and am immersed in a sea of data about it, my inability to place everything before me into a mutually consistent relation will be frustrating. I want the world to make sense…and that is what the conspiracy theory does, and does with enviable economy. Of course to a certain kind of outsider, such a thought process may appear to be some kind of cognitive infection, but this will only be so for those skeptics who already implement *another* rational process; specifically, the one concerned with identifying and disregarding non-falsifiable claims. That process may itself depend on yet others, and so on.

Thus, the kind of the fitness that induces *cultural* transmission is relative to the mental ecology of the individual. It is the same for the recipient: his or her cognitive ecology might be such that a particular process, well suited to the donor, has low fitness in the new environment and so soon disappears.

The reader will recognize here the concept of *memes*, i.e. culturally transmissible concepts and practices (Dawkins, 1976). The fruitfulness of memetics as a scientific enterprise has been challenged (e.g. Atran, 2001) and I will not have the space here to address all these criticisms adequately, so will make only a few remarks. First, one of the central complaints about memes is that, since they are ideas, and ideas are so easily changed, memes are rarely copied with sufficient fidelity for them to qualify as units of selection. My response to this complaint is that while epistemic reasoning-related processes are ideas in a general sense, they are atypical and few in number. These few are indeed variable—indeed they would have to be, to satisfy another of the conditions for evolution—but as a matter of fact they are copied with high fidelity: no one thinks that only *blue* objects are permanent, or that the denominator of the t-statistic ought to be the fiftieth power of the standard error. So this first criticism of memes doesn't stick. Reasoning-related processes are preserved with considerable fidelity, both through vertical (genetic) and horizontal (cultural) transmission. In fact, the very possibility of genetically heritable reasoning processes means that some of these processes get their

integrity from the accuracy of the DNA copying process, just like other phenotypic traits, and clearly that process is faithful enough to allow genes to evolve.

A second, related complaint about memes is that they are hard to individuate. Genes make proteins, and in general half of a gene is no longer a gene because the protein it makes has no useful role. But memes are ideas, and half of an idea is still an idea. Thus anything that can be thought or communicated is a meme, and the concept is empty because undiscriminating. My answer to this is that reasoning processes appear to be at least somewhat discriminable components of cognitive machinery—witness the wide variety of challenges that cognitive psychologists have devised to expose them and the developmental and neural accounts of their ontogeny. Like genes, they cannot be subdivided without destroying function: half of a belief in modus ponens is no more functional than half a hemoglobin. So if the processes I am talking about are memes, then they are pretty individuated ones.

Let me pause here for some wayfinding. So far I have defined key terms (functional vs. structural distinction); described a quantitative method for comparing reasoning-related functions (Turing encodings and the edit distance) and put these into a genealogical tree; and explained how a person's variability in reasoning-related function constitutes a traversal of that tree. In doing this I have implicitly made cases for P1 (some reasoning-related processes are structurally distinct), P2 (people are diverse with respect to these processes), and P3 (movement through the tree is profitably seen as an evolutionary process). Now I can discuss how and why species concepts, repurposed from their original biological context, are helpful in understanding the types of differences and changes just reviewed.

4. Translating Applicable Species Concepts

There are more than twenty species concepts in biology but I will use just three: biological, phylogenetic, and ecological. In brief, the way I propose to repurpose these concepts is by distilling them to their foundations in interactions, origins, and phenotypes, respectively. Before I get to the repurposing, however, I want to situate my approach relative to a few landmark issues and authors in the species concept literature.

In my account, a person's style of epistemic reasoning is determined entirely by intrinsic structural properties; i.e. those of the particular Turing machine that their brain instantiates. One might think that for this reason, I am committed to an essentialist view of species, i.e. an assertion that some definite set of intrinsic properties constitutes a necessary and sufficient condition for species membership. Such a view is defended for biological species by, for example, Devitt (2010). Although I agree with Devitt on many points, I am not exclusively committed to essentialism. Neither do I reject it. I am pluralistic: I think that there is more than one sound way of grounding taxonomic concepts.

It must be understood at the outset that essentialism is not a popular view amongst philosophers of biology. Instead, they tend to prefer *relational* and especially *historical* theories in which species membership depends not on an organism's having a particular genetic constitution, but rather on its standing in some relation to other organisms. Part of the reason for this preference is (what I understand to be) the empirical fact that species-unique gene sets have not been found; another reason is that some criteria used in

marking species boundaries, such as reproductive isolation, are clearly contingent not intrinsic properties of individual organisms. Then, of course, there is the fact that species evolve, making it difficult to anchor a species on a set of fixed genetic properties. Devitt's defense hinges on a distinction between what he calls the *taxon* and the *category* questions. The taxon question is "In virtue of what is an organism a member of a certain species?" and the category question is "In virtue of what is a group of organisms a species?" Devitt's target is the taxon question, and his answer is that two organisms are members of the same species in virtue of their having particular structural (typically genetic) properties. Even when such properties are superficially relational (such as the property of being able to breed with a conspecific) they ultimately unpack into structural properties pertaining to the individual. Historical properties are clearly important for species, but they pertain to the category question, not the taxon question. Devitt concludes that a conflation between the two questions has confused other authors.

Ereshefsky (2010), however, rebuts this by asking *why it is the case* that organisms have the structural properties that they have. His answer: because they are genealogically connected to other organisms in a single lineage. Thus history is explanatorily prior to structure. Indeed, we could hypothetically change those structural properties in an individual without voiding its membership in the species. Clearly then, for Ereshefsky, it is in virtue of an historical property, not an intrinsic structural one, that an organism belongs to a certain species.

To resolve this dispute, albeit in a Solomonesque fashion, we might turn to Kitcher's (1984), which recognizes more than one valid species concept, motivated by a distinction between "structural" and "historical" explanations of organismal properties. On my reading, Kitcher's distinction—also used by Devitt—is just a partial recasting of Tinbergen's (1963) famous framing of four kinds of explanations in biology—a framing itself inspired by Aristotle's four causes. To wit, these four questions concern (1) causation, i.e. the mechanism of function; (2) ontogeny, i.e. the development of that mechanism in the life of the individual; (3) adaptation, i.e. the evolutionary pressures giving rise to the feature; and (4) phylogeny, i.e. the historical relations of that feature to extant others. Each of these questions has motivated some of my discussion of human reasoning, and I expect each will be emphasized differently by systematists—whether in biology or psychology—with different purposes.

To illuminate the value of a pluralistic approach, consider that for some purposes, the historical process through which a person came to instantiate a particular Turing machine does not matter. For example, if one were interested in normative assessment or interpersonal compatibility, present function would be all-important, and species boundaries could be drawn entirely ahistorically. But suppose one were interested explicitly in the development, within an individual, of epistemic reasoning, or in education. I said earlier that not all trajectories through the Turing tree are equally likely. So it could be that two people close to each other in reasoning space (per the edit distance on their Turing encodings) are likely to have divergent futures. In such a case it might be more fruitful to base species membership not on absolute position in Turing space, or even on difference from each other, but instead on position relative to attractors in that space—which is to say, local optima of the fitness landscape.

In biology, John Dupré (1981, 2012) favors an attitude even more inclusive than Kitcher's, which he calls *promiscuous realism*. By this he means that even non-biologists

have classificatory needs to which species concepts can be validly applied. There are differences between people, and the differences are real, but the classificatory move itself is purely pragmatic. A chef will have need of one classificatory system for fungi, a mold-removal specialist another, and a physician yet another; perhaps something similar will hold for neuroscientists, developmental psychologists, propagandists. This is largely why I will introduce three ways of thinking about psychological species instead of just one: I think that species concepts are powerful, broadly applicable, and I am relatively unconcerned with whether the classificatory boundaries they might make in the space of epistemic reasoning constitute cuts at nature's joints.

There are, as it turns out, conceptualizations of species that do justice to structural and historical explanations, both at once, and I find these both appealing and powerful. For example, Boyd (1999) articulates a view of species as natural kinds with essential properties, but instead of assuming that essential properties must be lists of biological (e.g. genetic, physiological, or morphological) traits, Boyd takes them to be sets of *any* properties, broadly conceived, that confer a collective homeostatic tendency. That is, a species' essence is a set of properties that tends to result in the stability over time of that set. Such properties are not restricted to biological traits but could include extrinsic properties, including but not limited to the example of reproductive isolation given earlier. Moreover, the properties most significant to homeostasis may change over time as the population and the environment within which it is embedded exert variable pressures. Boyd calls this approach Homeostatic Property Cluster (HPC) theory. The way HPC theory would be applied here would be, roughly, to map the needs for survival and personal sense-making onto a set of properties—not just substrings of Turing encodings, but perhaps circumstantial properties like habitual proximity to and engagement with a text or texts—that met these needs. In this way each psychological species would be like an island of mutually-reinforcing epistemic processes, internally consistent and able to deal with new experiences but incompatible with the functional ecologies of other islands. Although I did not develop the ideas here with HPC theory in mind, and I don't share all of Boyd's aims, I am optimistic that such an application could work—though I won't attempt it here.

Now to the details. The first species concept that I want to explore is the most well-known: the biological or reproductive concept. Reproductive conspecifics are members who can and do interbreed to produce fertile offspring. Translating this into the reasoning domain requires a psychological analog for reproductive compatibility. I am modeling a person's reasoning-related psychology as a set of processes, so one of the tools needed here is a pairwise hybridization procedure like meiosis, i.e. one that produces a new set with some (50%, let's say) of its processes drawn from one parent and the rest from the other. A person implementing the processes associated with that hybrid set would be the psychological offspring of the two parents. By itself, this is not enough, because such hybrid sets can always be produced, no matter what the parent sets. We need to determine when the offspring are "fertile" in a way that goes beyond the ability to donate material to another daughter set.

What I suggest is this: a reasoning-related system that generates internally contradictory beliefs should be regarded as infertile.

Consider what happens when two patient and intellectually honest people engage in analytical discourse, trying to determine the compatibility of their beliefs and the

methods by which they arrive at them. In ideal circumstances, this conversation provides opportunities for each party charitably to entertain the suppositions and methods of the other. These entertainments—person A suspends their own opinions temporarily to assume B's point of view—can be seen as temporary hybrid beings, part one speaker and part the other. There need not be any actual being who inherits processes from both parties: we need only see what happens to a *hypothetical* person with that inheritance. If that hybrid can work through a collection of reasoning problems without self-contradiction, then the two original parties would register as conspecifics under this version of the species concept.

The central idea here is hybrid compatibility. In biological reproduction, the hybrid must demonstrate intact functioning of its reproductive machinery—machinery that works only if a vast dependency diagram of metabolic processes is unperturbed. With reasoning-related processes, intact functioning is also required, but here the dependency diagram is based on compatible patterns of mental and neural representations rather than chromosome counts or DNA reading frames. Compatibility of this sort depends on processes but not, in general, on the content of the doxastic "database." That is, it is possible for two people who are "reproductively" compatible under this relation, and therefore members of the same psychological species, to have contradictory beliefs. This may arise because the people were exposed to different content, even though the methods by which they derive beliefs from that content are compatible. The hybrid, though, would have its own experiences upon which its reasoning processes operate, and if that could happen without contradiction, then the two parents would be conspecific. It is at that deeper level—*how* we believe, not what we believe—that we ought to base our classificatory discriminations.

The combinatorics of hybridization means that with a large set of processes, there are very many possible offspring. Presumably some processes are more important than others, so it is possible that some hybrids will generate internal contradictions and others will not, or that a given hybrid will generate contradictions on certain types of problems but not others. So although conspecificity appears to be an all-or-none judgment, it is actually fuzzier than that. It is much the same in biological species, and the interoperability of the individuals involved can depend on the particulars of which functions bear the difference. Mules (the offspring of a female horse and a male donkey) are normally infertile because their parents have unequal numbers of chromosomes. For this reason horses and donkeys are generally considered different biological species. Occasionally, however, fertile mules are produced, apparently because the sex chromosomes are entirely from one parent while the autosomes are hybrids. An analog in the psychological realm might result from the interpersonal compatibility of some critically important processes (the one that determines, for instance, whether an assertion is compatible with other well-established beliefs) and incompatibility of comparatively less important ones (such as the one that determines whether an assertion has been properly vetted by peer review). Determining which processes are critical and which are not would require a model of cognitive architecture more sophisticated than anything I'm prepared to advance here.

The second species concept I want to explore is the phylogenetic species concept, under which taxa are based on common ancestry rather than hybrid compatibility. Thus this concept counts origin as more relevant than present function. To apply this to the

context of this paper, consider the lineage and variants of the reasoning-related processes used in the scientific method itself. This style of inquiry has a rich and well-documented ancestry that I needn't review, but note that it is not just one method but many. In sciences yoked closely to deductive-nomological practice (e.g. physics and perhaps economics), belief is normally withheld from any assertion that cannot demonstrate consistency with a framework of well-established quantitative laws—or, at least, confidence falls somewhat if a proposed new law cannot be derived from a such a framework. In more inductive disciplines (e.g. anthropology, or, for that matter, psychology), there are few or no widely agreed-upon broadly applicable theoretic models or quantitative laws, and not much in the way of deductive-nomological derivation. Belief then derives from an abductive process based on a perceived preponderance of the evidence in comparatively small subdisciplines.

Despite their present-day differences, these fields, and the sets of reasoning-related processes implemented in the minds of their practitioners, both owe their existence to the Aristotelian traditions of logic and empiricism. Because of that shared ancestry they are in one sense more meaningfully related to each other than either is to, say, comparative literature. Reasoning-related diversity reaches even into the "pure" sciences: Mathematics itself may be even now in the process of speciating around the admissibility of computer-assisted and probabilistic proofs (Fallis, 2011; McEvoy, 2008). Some physicists see string theory, which as far as I know has struggled to produce empirically testable claims, not just as *non-physics*, but as some kind of theology. Again, to construct an evolutionary taxonomy of these traditions, we would want a Turing encoding of all the processes used in their normative practice. Then we would compute edit distances between every pair of encodings and infer from those distances a lineage: the smaller the distance, the more recent the common ancestor.

The same sort of taxonomy can be found outside of science. The history of the Christian church illustrates the spatiotemporal spread and functional modification of doxastic processes. The early Christian canon hybridized with local pagan practices differently in different places, suffered the heresies of Martin Luther, John Calvin, and Joseph Smith, and settled into the minds of countless millions. Many of these people, in an effort to square their inheritance with intuition and experience, i.e. to make sense of things, have grafted and elided their way into single-serving belief systems. At one time, powerful selective forces would have extinguished these mutants, but in the Western world at least, many of those forces have faded. Indeed the "religious but nothing in particular" are a growing group (Pew Forum on Religion and Public Life, 2012).

In all these cases we find clusters of practices concerning how to decide what to believe. These practices compete on a fitness landscape (though in these cases the fitness is perhaps more of the sense-making kind, not the Darwinian kind) and rise or fall based on the outcome of that contest. Clearly they change over time, and are transmissible from generation to generation. These are the evolutionary trees of psychological species.

The third and last species concept that I want to repurpose is the ecological species concept, which in biology is based on exploitation of a single niche: organisms are ecologically conspecific if their survival is based on using the same resources in the same way. To do this, they generally must share morphological and metabolic features. In the present context, such features correspond to reasoning-related processes, so a group of people (or the processes they implement) are conspecific if they are weakly equivalent,

irrespective of whether this equivalence is due to independent discovery or common descent.

As an illustration of this last type of psychological species, consider the process I will call *titration.* Your problem is to determine how much of some intervention is required in order to obtain a desired result. The titration method is to gradually administer small amounts of the intervention until the target is reached. This is not the only way to solve the problem: you could consult reference material, make an inference based on similar experiences, or guess. Nor is it necessarily the best: there could be time pressure or undesired consequences of gradual administration (shall we extract the tooth with a slow and steady pull or a sharp jerk?). A large and disjoint set of people have discovered and applied titrimetry to a broad range of problems: Chefs might use it to determine how much salt to add. Cement mixers use it to determine how much water, retailers how much to discount, dog trainers how many treats, and romantic suitors how many gifts. The convergence of these groups on a single means of problem solving is due neither to hybridizing interactions between individuals, nor to inheritance from an ancestral practice, but rather to independent discovery. The processes implemented by these people share a structure, in this detail at least, because the problems they confront are amenable to the application of the titration method. This is also how it is with ecological species. The organisms we classify as ecological conspecifics are generally also the same phenetically (i.e. they have the same structure) because the niche they exploit imposes certain physical requirements: a beak of a certain shape is required to extract nectar from flowers. The people in the aforementioned groups are psychologically and behaviorally the same because the problem they face imposes similar constraints.

Another example: consider that many religions evolved independently the practice of determining truth by consulting a privileged individual or role (the shaman, the priest, the Pope). Presumably this practice arose out of a desire to maintain doctrinal coherence, a power structure, etc. That common use of the *belief-from-authority* process places them in the same (psycho-)ecological species with each other, and in a different species from scientists, whose constraints are such that registration with empirical results (among other things) dominates over preservation of a power structure or doctrinal survival. But some of these religions appeared on different continents, thousands of years apart. I suppose it is possible that their similarity is due to inheritance of genes governing primate societies, but I rather suspect that it is instead that they discovered the solution independently.

To conclude this section, I want to defend my case against a possible objection, which is that by extending species concepts to encompass abstract functional differences, the concepts become so general as to be worthless. Could we not say, using this scheme, that *any* two distinct things—two grocery store designs, say—were members of different species? Presumably the objective properties of any given store—aisle width, size of the produce section, etc.—could be quantified and compared with those of another. We could combine the features of two stores to create a hybrid. If species concepts can be applied so universally, the objection goes, then they lose their meaning.

My response is as follows. I do not propose that species concepts can be usefully applied to any two distinct things. Although a rock and an electron are functionally different, I would agree with my objector that it would be fruitless to ply species concepts to that difference. I *would*, however, hold out some hope that these concepts could be

applied to grocery stores, and the reason is that grocery stores, like living things and psychologies, are *evolved designs*. Grocery stores are designed and modified over time by intelligent agents while living things and psychologies are designed by the forces of natural evolution. To regard species concepts as being applicable only to biology is to miss much of their explanatory power. These concepts provide sciences other than biology a powerful toolkit that helps make sense of the histories and relationships of evolved designs, *tout court* (Cziko, 1997). All such entities are subject to selective forces, at the mercy of which they either persist or disappear. They all compete on some kind of fitness landscape. They all mutate, either by chance or goal-directed intention. Rocks and electrons are not subject to these forces, so even though they can be computationally characterized, they are not evolved designs and species concepts ought not be applied to them.

5. Other Classificatory Systems in Psychology

Folk psychology has long marked boundaries in personality space with distinctions like sanguine and phlegmatic, hot and cold blooded tempers, the sane and the mad. In this section, I'll review a few of these classificatory systems and consider whether any is well-suited to an application of a repurposed species concept.

The longest-lived system of psychological classification may be Galen's four temperaments theory based on the Aristotelian elements, while the most well known modern theory is probably the Meyers-Briggs Type Indicator (MBTI), which flows from Jungian personality theory. Because both of these systems discretize psychological space, they are referred to as *type* models. More recently, *trait* models, which allow continuous variability along the axes of personality, have displaced the type models. The most well known trait theory is the Five Factor Model (FFM), which rates people according to their openness to experience, conscientiousness, extraversion, agreeableness, and neuroticism.

A central question for any theory of personality is which, and how many, traits to include, and it is here that we see a difference from theories of reasoning. In the latter case, the processes that matter can be found out because beliefs have representational content: their role in psychology is to capture the way things are. And very often, the way things are can be determined by objective empirical means. So if by turning the dial on a process we can lead someone into a demonstrably false belief, then that process ought to be included in a theory of epistemic reasoning. The FFM trait extraversion is indubitably a psychological process, but no matter how we turn the dial on its parameters, we cannot lead someone into a false belief—we can only make them like or dislike parties.

It is not clear that either the FFM or the MBTI has a comparable means of selecting the processes that matter for personality. The Big Five traits are apropos for many academic psychologists in western industrialized cultures but there are other traits not included in the FFM that might explain other sorts of psycho-behavioral regularities. For example, there is nothing in the FFM about respect for tradition. This may be because western industrialized cultures are generally not heavily invested in tradition. Some others are, however: Cheung & Leung (1998) found that respect for tradition helped explain variability on personality instruments administered to Chinese subjects. Going in the other direction, it seems plausible that neuroticism (for instance) might tend to be more salient in professional, largely anonymous urban environments than in agrarian

small communities, and indeed Gurven, von Rueden, Massenkoff, Kaplan, & Vie (2013) argue that only two, not five, components of personality explain most of the behavioral variance among farmers in rural Bolivia. Environments create spaces for certain psychological tendencies but not others, and psychologists tend to be interested in the tendencies dominant in the culture to which they belong.

My point here is that the criteria for what constitute relevant features of personality vary with culture, and this makes them a less suitable basis for species-like classifications than the processes of epistemic reasoning.

Let me turn now to another psychological classificatory system, this time one that can better substantiate its claims about which traits matter for interpersonal difference. Jonathan Haidt's *moral foundations theory* aims to explain the apparently irreconcilable moral / political views of strongly liberal and strongly conservative people (Haidt & Graham, 2007; Haidt & Joseph, 2004). To do this, it claims that the moral reasoning process for both groups is a mechanism that evaluates situations along five dimensions: fairness/reciprocity, harm/care, ingroup/loyalty, authority/respect, and purity/sanctity. The difference between liberals and conservatives arises because liberals weight the first two of these dimensions, i.e. fairness and harm, much more highly than the other three, while conservatives weight all five components roughly equally (Graham, Haidt, & Noseck, 2009).

On the view I am advocating here, moral foundations theory by itself does not press us toward a species-like classification of liberals and conservatives—despite the frequent appearance of irreconcilability between these groups—for two, perhaps three reasons. First, people at opposite ends of the liberal/conservative spectrum agree, according to the model, about what ought to figure into a moral calculation. Even strongly liberal people agree with strongly conservative people that ingroup, authority and purity factors matter in moral decisions. Thus the two groups have compatible moral decision-making processes: unlike the single- and multi-layer perceptrons, the networks at work in liberals and conservatives differ only parametrically. Second, although Graham et al. talk in terms of "liberals and conservatives," these do not appear to be separable clusters on the liberalism/conservatism axis. The more conservative a person is, the more uniform the weighting of the five factors: the actual data vary smoothly from one end of the axis to the other.

There is yet a third reason to resist the application of species concepts in the moral space explored by Haidt and colleagues, and this applies to personality psychology as well. One of my motivations for focusing on reasoning-related traits is that these traits have more profound consequences for our social life—and indeed, for long-term global flourishing—than those typically included in personality psychology or moral / political reasoning. This is not to say that there are no possible grounds for drawing species boundaries along these latter or other axes. It may be, for all I know, that people in the 80th percentile on the extraversion trait are structurally dissimilar in some way from people in the 20th percentile. To make that determination, we would need a theory of personality that provided functional and perhaps neural mechanisms for that trait. If we had such a theory, and could show that high- and low-extraversion scorers were thereby functionally incompatible, we ought to be prepared to say that they were members of different species with respect to those functions.

Although moral foundations theory doesn't at present warrant the drawing of species boundaries in the space of moral and political thinking, some other theory of political decision making might. Still, I maintain that the psychological mechanisms by which a person decides what is true (as opposed to what is preferable, or socially relevant) are more general, more fundamental, and have more profound consequences than the mechanisms that determine either their personality traits or their political affiliations. And they are certainly more consequential than the criteria sometimes used to demarcate biological species.

6. Causes of Psychological Speciation

Most people agree about ordinary things: whether there is a cup on the table, whether it is raining, etc. Our shared world, shared needs, and rich communication about these things have conspired to ensure doxastic compatibility across a wide range of mundane propositions, leaving room for real structural divergence mostly on matters where the answer is not evident to direct observation. It follows that releasing either of two constraints—a shared world or frequent communication—can expedite functional divergence. In other words, psychological speciation is facilitated when (1) ground truth is not readily accessible, and/or (2) people are not disposed to coordinate their reasoning processes.

Situations of the first type arise perhaps most commonly in religion and other metaphysical domains, but this is not exclusive. In politics, for instance, the complexity of a large society and the difficulty of doing controlled experiments means that it can be difficult to determine when social policy or program is, was, or would be effective. Even if two people agree completely about what constitutes a socially desirable outcome, they may, by attending to and interpreting data differently, come to very different conclusions about the best way of achieving it. This may point toward an alternative to Haidt's moral foundations theory of political difference: perhaps it is not just, or not primarily, differing moral commitments that matters, but differing epistemic processes. On that view, the radical species-like difference between a Marxist and a capitalist could be a consequence of the difficulty in determining how economies work, the fact that brute facts about them are hard to come by. C. P. Snow's two cultures—a classic example of group difference—may have diverged for equally epistemic reasons. The facts of literature or art or rhetoric are more slippery things than the facts of natural and physical sciences, so it is easier for two art theorists to find themselves at loggerheads than two mathematicians.

In addition to requiring access to unambiguous truths, the maintenance of interpersonal rational compatibility depends on a steady habit of talking about those truths and their implications. This is the second constraint mentioned above. In biological evolution, the prolonged isolation of a splinter group from its root population can lead to speciation, both because unique selective pressures act on the group and also because of founder effects and genetic drift. Isolation in those biological cases is often caused by physical obstacles. Psychological isolation can be brought about in the same way, but it can also be the result of *functional* rather than physical barriers such as social taboos regarding certain topics or emotional factors causing us to avoid contrary people or information. Analytical discourse—the effort to make two reasoning machines work together harmoniously—can thus be seen, as I said earlier, as a kind of psychological

analog for sexual reproduction. The slow divergence of beliefs brought about by polite silences and parochial schools facilitates psychological speciation through analogs of selection and genetic drift.

I take it as an uncontroversial assertion that there are developmental influences on reasoning-related habits, so this too could be a source of functional diversity. A child subject to positive or negative reinforcements for the exercise of particular practices may be more or less disposed to carry those practices into adulthood. Parents and teachers will use such reinforcements, both consciously and unconsciously, to instill their particular brand of rationality in their charges. Beyond such deliberate direction, we might wonder about more implicit influences of the developmental landscape. Language is a controversial possibility, figuring into cognitive diversity via the thesis of linguistic relativity (alias Whorfianism). I make no assertions about the relationship between reasoning processes and native speaking language or even any putative language of thought (Aydede, 2010), though I note in passing that some recent results seem to modestly contraindicate such an association: Perner, Mauer, & Hildenbrand (2011) found that the sophistication of children's reasoning processes were not strongly correlated with their general linguistic competence. Bearing perhaps more interestingly on this issue is the existence of spoken languages whose grammar requires modifiers indicating, for instance, whether one came by a belief through direct observation, hearsay, or inference (Aikhenvald, 2006; Fleck, 2007). Future research could reveal whether native speakers of such languages are perhaps more robust reasoners.

As a final possible etiology of reasoning-related difference, consider mental illness, in particular the famous Capgras delusion, in which a patient believes that someone, typically a family member, has been abducted and replaced with an impostor or perfect robotic replica. As bizarre as these beliefs are, these are sincerely held and actionable: in one tragic example, a patient decapitated one of the "replica" family members in order to find the batteries and microfilm inside (de Pauw & Szulecka, 1988). Most people would readily classify Capgras patients as psychotic, and thereby agree with me on the core contention that there are real and important boundaries to mark out in doxastic space. But there is a surprise in store here. While I readily agree that some mental illnesses might banish a person into a rather remote corner of that space, Capgras in particular might not be an unequivocal example of doxastic dysfunction. Stone & Young (1997) explain why.

First, they say, Capgras patients suffer from a *perceptual deficit* in which the affective information attached to a face becomes unavailable. The patients can still recognize friends and family, but their phenomenal experience is stripped of emotional associations. Second, Capgras patients display a *reasoning bias*, not a dysfunction, such that they weight perceptual data more strongly than consistency with established beliefs. The combination of these two factors yields a line of thought, for the Capgras patient, that goes like this: "I see that this person looks exactly like my mother, but if she were actually my mother, I would feel something when I look at her. I do not feel anything, therefore she cannot be my mother." This is a perfectly intelligible, even laudably normative rational inference. Of course, a typical subject in these circumstances might go on to think, "It is implausible, based on my understanding of human biology etc., that there could be an exact replica of my mother, so something must have gone wrong with

my perceptual system." But Capgras patients are biased in favor of their perceptual system so do not do this.

By these lights, Capgras patients possess and exercise rational processes that are compatible with those of normal subjects. The difference between them and normal people is values-based, not reasoning-based. "Accounts in terms of biases thus reduce the distance between delusional and normal beliefs, seeing them as on a continuum rather than in separate categories to be labeled as 'rational' or 'clearly mad'" (Stone & Young, 1997, p. 343).

Before closing, I would like to say one more thing about the apparent disanalogy between biological and psychological species, namely the durability (or lack thereof) of an individual's species membership. Biologically, individuals do not change species, but I have argued that psychological species changes are the norm, not the exception. Observe, however, that there is no really principled reason for the fixity of biological species membership. A freakish burst of cosmic rays or chemical mutagens could induce such a switch—at least, if you think that structural properties fix species membership. It just happens that the probability of that happening is exceedingly low. Dupré (1981, p. 88) amusingly argued that "if a chicken began to lay perfectly ordinary walnuts which were planted and grew into walnut trees, [he] would not wish to refer to this result as the production of a grove of chickens." But because neuronal networks are so much more plastic than DNA sequences, it is quite possible that a person could make such an inter-species move. Indeed most people probably do so in the course of normal development and education (Santamaría et al., 2013). More rarely, organic brain disease could induce a species switch, and yet another possibility is supplied by Pyysiäinen (2003, pp. 119,121), who offers an account of abrupt doxastic change in adults when they undergo religious conversion experiences.

7. Conclusion

Imagine looking from a great height, as it were, at this tree of rational processes, each node of which represents a distinct way of thinking. Now place every person alive on their respective node. Most nodes will be empty. Others will be occupied by just a few people, and others will teem with thousands and millions. Now turn back the clock and witness the migration of our species across this landscape: a crowd gathers on a particular spot and waits there, stuck, until a bridge appears to another node, another way of thinking. Someone discovers something, migrates to new territory alone, then beckons for others to join with a package, wrapped in language, containing the computational process that got them there. These migrations are driven by gains in two types of fitness: classical biological fitness and the positive affective response that comes from individualized sense-making. Although the interpersonal differences that result from this evolutionary process are invisible, they are as real as the morphological differences between species and critically important because of the pivotal role that epistemic reasoning plays in guiding human behavior.

References

Aikhenvald, A. Y. (2006). Evidentiality in grammar. In K. Brown (Ed.), *Encyclopedia of Language and Linguistics* (2nd ed., pp. 320–325). Oxford: Elsevier.

Atran, S. (2001). The trouble with memes: inference versus imitation in cultural creation. *Human Nature*, *12*(4), 351–381.

Aydede, M. (2010). The language of thought hypothesis. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2010.). Retrieved from http://plato.stanford.edu/archives/fall2010/entries/language-thought/

Baird, A. (2002). Frontal Lobe Activation during Object Permanence: Data from Near-Infrared Spectroscopy. *NeuroImage*, *16*(4), 1120–1126. doi:10.1006/nimg.2002.1170

Boyd, R. (1999). Homeostasis, species, and higher taxa. In R. A. Wilson (Ed.), *Species: New Interdisciplinary Essays* (pp. 141–185). Cambridge, MA: MIT Press.

Buzsáki, G. (2010). Neural syntax: cell assemblies, synapsembles and readers. *Neuron*, *68*(3), 362–385.

Cheung, F. M., & Leung, K. (1998). Indigenous personality measures: Chinese examples. *Journal of Cross-Cultural Psychology*, *29*, 233–248.

Corradini, A., & Antonietti, A. (2013). Mirror neurons and their function in cognitively understood empathy. *Consciousness and Cognition*, *22*(3), 1152–1161. doi:10.1016/j.concog.2013.03.003

Cziko, G. (1997). *Without miracles universal selection theory and the second darwinian revolution*. Cambridge, Mass.: MIT Press.

Dawkins, R. (1976). *The Selfish Gene*. Oxford: Oxford University Press.

De Pauw, K. W., & Szulecka, T. K. (1988). Dangerous delusions: violence and the misidentification syndromes. *British Journal of Psychiatry*, *152*, 91–97.

Devitt, M. (2010). Species have (partly) intrinsic essences. *Philosophy of Science*, *77*, 648–661.

Dupré, J. (1981). Natural kinds and biological taxa. *Philosophical Review*, *90*, 66–90.

Dupré, J. (2012). *Processes of Life: Essays in the Philosophy of Biology*. Oxford University Press.

Ereshefsky, M. (2010). Species. In E. N. Zalta (Ed.), *Stanford Encyclopedia of Philosophy* (Spring 2010.). Retrieved from http://plato.stanford.edu/archives/spr2010/entries/species/

Essex, B. G., Clinton, S. A., Wonderley, L. R., & Zald, D. H. (2012). The Impact of the Posterior Parietal and Dorsolateral Prefrontal Cortices on the Optimization of Long-Term versus Immediate Value. *Journal of Neuroscience*, *32*(44), 15403–15413. doi:10.1523/JNEUROSCI.6106-11.2012

Fallis, D. (2011). Probabilistic proofs and the collective epistemic goals of mathematicians. In H. B. Schmid, M. Weber, & D. Sirtes (Eds.), *Collective Epistemology*. Ontos Verlag.

Fleck, D. (2007). Evidentiality and double tense in Matses. *Language*, *83*(3), 589–614.

Godfrey-Smith, P. (2002). Environmental complexity and the evolution of cognition. In R. Sternberg & J. Kaufman (Eds.), *The Evolution of Intelligence* (pp. 233–249). Mahwah, NJ: Lawrence Erlbuam.

Graham, J., Haidt, J., & Noseck, B. A. (2009). Liberals and conservatives rely on different sets of moral foundations. *Journal of Personality and Social Psychology*, *96*(5), 1029–1046.

Gurven, M., von Rueden, C., Massenkoff, M., Kaplan, H., & Vie, M. L. (2013). How universal is the Big Five? Testing the Five-Factor Model of personality variation among forager-farmers in the Bolivian Amazon. *Journal of Personality and Social Psychology*, *104*(2), 354–370.

Haidt, J., & Graham, J. (2007). When morality opposes justice: Conservatives have moral intuitions that liberals may not recognize. *Social Justice Research*, *20*, 98–116.

Haidt, J., & Joseph, C. (2004). Intuitive ethics: How innately prepared intuitions generate culturally variable virtues. *Daedalus: Special Issue on Human Nature*, *133*(4), 55–66.

Kitcher, P. (1984). Species. *Philosophy of Science*, *51*, 308–333.

Levenshtein, V. (1966). Binary codes capable of correcting deletions, insertions, and reversals. *Soviet Physics--Doklady 10*, *10*, 707–710.

Mata, A., Ferreira, M. B., & Sherman, S. J. (2013). The metacognitive advantage of deliberative thinkers: A dual-process perspective on overconfidence. *Journal of Personality and Social Psychology*, *105*(3), 353–373. doi:10.1037/a0033640

McEvoy, M. (2008). The epistemological status of computer-assisted proofs. *Philosophia Mathematica*, *16*(3), 374–387.

Minsky, M., & Papert, S. A. (1969). *Perceptrons*. Cambridge, Mass.: MIT Press.

Nei, M. (1987). *Molecular evolutionary genetics*. New York: Columbia University Press.

Perner, J., Mauer, M. C., & Hildenbrand, M. (2011). Identity: key to children's understanding of belief. *Science*, *333*, 474–477.

Pew Forum on Religion and Public Life. (2012). *"Nones" on the rise*. Retrieved from http://www.pewforum.org/Unaffiliated/nones-on-the-rise.aspx

Pylyshyn, Z. (1984). *Computation and cognition*. Cambridge, MA: MIT Press.

Pyysiäinen, I. (2003). True fiction: philosophy and psychology of religious belief. *Philosophical Psychology*, *16*(1), 109–125.

Reverberi, C., Bonatti, L. L., Frackowiak, R. S. J., Paulesu, E., Cherubini, P., & Macaluso, E. (2012). Large scale brain activations predict reasoning profiles. *NeuroImage*, *59*(2), 1752–1764. doi:10.1016/j.neuroimage.2011.08.027

Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*, *65*(6), 386–408.

Santamaría, C., Tse, P. P., Moreno-Ríos, S., & García-Madruga, J. A. (2013). Deductive reasoning and metalogical knowledge in preadolescence: A mental model appraisal. *Journal of Cognitive Psychology*, *25*(2), 192–200. doi:10.1080/20445911.2012.743988

Stanovich, K. E. (2012). On the distinction between rationality and intelligence: Implications for understanding individual differences in reasoning. In K. Holyoak & R. Morrison (Eds.), *The Oxford Handbook of Thinking and Reasoning* (pp. 343–365). New York: Oxford University Press.

Stone, T., & Young, A. W. (1997). Delusions and brain injury: the philosophy and psychology of belief. *Mind and Language*, *12*(3/4), 327–364.

Svedholm, A. M., & Lindeman, M. (2013). The separate roles of the reflective mind and involuntary inhibitory control in gatekeeping paranormal beliefs and the

underlying intuitive confusions: *Gatekeeping paranormal beliefs*. *British Journal of Psychology*, *104*(3), 303–319. doi:10.1111/j.2044-8295.2012.02118.x

Tinbergen, N. (1963). On aims and methods of Ethology. *Zeitschrift Für Tierpsychologie*, *20*, 410–433.

Turing, A. M. (1937). On Computable Numbers, with an Application to the Entscheidungsproblem. *Proceedings of the London Mathematical Society*, *s2-42*(1), 230–265. doi:10.1112/plms/s2-42.1.230