

## Hard Problems, Interpretive Concepts, and Humean Laws

*Apart from conscious experience and perhaps indexicality, these [facts about physical laws and causation] constitute the only such **further facts** [facts not explicable in terms of particular physical facts] in which we have any reason to believe. It is not unnatural to speculate that these two nonsupervenient kinds, consciousness and causation, may have a close metaphysical relation. (Chalmers 1996, p. 86)*

*An interpretation is **not** evidence of some **further fact**. A true interpretive claim is true because the reasons for accepting it are better than the reasons for accepting any rival interpretive claim. (Dworkin 2011, p. 154)*

John F. Halpin, Oakland University

(Old unpublished manuscript, 2009ish version, that just lost its long-term hosting service. I have added the Dworkin epigraph from his *Hedgehogs* book as I post this. The idea is that the facts about consciousness arise in virtue of the brain's own self-interpretation, an underappreciated proposal in some cognitive science of consciousness in the last decades. This self-interpretation proposal is found prominently in contemporary Bayesian-brain and Predictive Processing accounts of the brain's generative self-modeling.)

Conceptual problems for consciousness are analogous to the Humean's problem with scientific laws. Just as consciousness appears to involve further facts beyond the physical, scientific laws may seem to involve a reality beyond the Humean's occurrent facts<sup>1</sup>. David Chalmers noted this long ago as quoted in the epigraph. Still, I will attempt to show that a Lewis-style best-system account for laws/scientific principles, one forswearing further facts beyond the occurrent, may be applied to the related problem for consciousness. The leading idea of a best-system account is that law and associated chance claims are true in virtue of their place in ideal systematic treatment of the totality of occurrent fact. Nomic facts about law and chance, then, are not "further facts" beyond the occurrent but only a matter of what makes for idealized scientific theory (or, if one prefers, idealized scientific interpretation).

There are reasons to prefer a "perspectival" version of the best-system account of law. For example, as I will argue, natural necessity can make sense to a Humean taking a practical perspective: It is practical to project *actual* laws (i.e., actual best-systems as best we know them) onto *possible*

---

<sup>1</sup> The Humean of this debate postulates only facts about mere happenings, so called "occurrent" facts independent of any potentiality.

worlds (thus taking the perspective of the actual world and imposing it onto other worlds). Also, different sciences have different perspectives for systematizing the actual occurrent facts. In this paper, I will try to show that a personal perspective *can* be best for systematizing psychological/neural occurrence. If this is right, then there is reason to believe that the so-called “further facts” of consciousness are just a matter of idealized systematization or interpretation from a personal perspective on the occurrences (specifically on occurrences regarding brain information-flow). Then, just as chance is vindicated in virtue of its place in ideal physics systematization, consciousness is vindicated in virtue of its place in the ideal theory of mind (from the brain’s native perspective or self-interpretation). Then, following the Dworkin quote of the epigraph, this essay concludes that our natural understanding of ourselves *as conscious* is not indicative of further facts. It is merely the brain’s self-interpretation, a native and evolved overlay to the more basic neural occurrences. Because David Chalmers first noted the law of nature-consciousness analogy and because he drew a contrary conclusion, I start with his well-known zombie argument.

## I

It would *seem* conceivable and even remotely possible that a being could behave precisely like you or me *without* having any conscious experience. That is, one can coherently imagine a “zombie” version of oneself. By definition, one’s zombie twin is physically and so behaviorally identical to oneself, yet has *no* inner mental feelings, no states with phenomenal character, no qualitative experience, no conscious point-of-view, and so forth. Most intuitions agree that such a zombie is conceivable even if not *nomically* possible.

David Chalmers, following in a long tradition running from Descartes<sup>2</sup> to Kripke<sup>3</sup>, argues that zombies of the sort just described are ideally conceivable and so<sup>4</sup> *metaphysically* possible. If a zombie physically indistinguishable from me is truly possible, then I differ from him only in my

---

<sup>2</sup>Descartes would presumably allow that non-human animals are zombies. But humans capable of language use must have conscious minds. Descartes sees us as differing from animals in virtue of a non-physical add-on to our brain. Famously, he argues that a mind without anything physical is conceivable.

<sup>3</sup>Kripke provides a number of these modal intuitions including a version of the zombie one; he writes “*Prima facie*, it would seem that it is at least logically possible that [brain state] B should have existed (Jones’s brain could have been in exactly that state at the time in question) without Jones feeling any pain at all, and thus without the presence of [pain sensation] A.” (1980, p. 146)

<sup>4</sup>The argument from conceivability (in Chalmers’ sense) to metaphysical possibility is an interesting and difficult one. My concern, however, will only be with the first step in the argument just given, the claim about conceivability.

conscious states. Because he and I are *physically* identical, it would seem that our difference, consciousness, must be of non-physical nature. This is a *momentous* conclusion. If the reasoning is sound, we have not only a proof of dualism, but one based on an apparently innocent and highly intuitive premise. Of course, given that the dualist conclusion is based on such a minimal basis, one must wonder about the reliability of the intuition and reasoning involved. Much concern in the literature has addressed the reasoning; I will focus on the zombie intuition itself.

It is worth noting that a Cartesian would *not* be concerned by the reliance on intuition for such a strong conclusion. After all, Descartes had God on his side to provide the guaranteed truth of clear and distinct ideas like the conceivability claim. On the other hand, Kripke (1980) does not reach a conclusion as strong as Chalmers'. Kripke utilizes zombie (and other modal) intuitions to get to the weaker conclusion that mental states/tokens are *not* identical to brain states/tokens. This is neither enough to deny all types of physicalism nor force one to the dualism espoused by Chalmers. Indeed, Kripke concludes *Naming and Necessity* with these words:

I suspect...that the present considerations tell heavily against the *usual* forms of materialism. Materialism, I think, must hold that a physical description of the world is a complete description of it, that any mental facts are 'ontologically dependent' on physical facts in the straightforward sense of following from them by necessity. No *identity theorist* seems to me to have made a convincing argument against the intuitive view that this is not the case. (1980, 155, *emphasis added*)

Still, about defenses of the identity theory, Kripke writes "[some] strike me as highly compelling arguments which I am at present unable to answer convincingly", and "I regard the mind-body problem as wide open and extremely confusing." (1980, 155, footnote 77) Thus, Kripke leaves open the possibility that a new sort of materialism may be more successful than the identity theory.

Central to my proposed materialism is the analogy between laws and nature and consciousness. Indeed, Chalmers addressed this analogy years ago from the antireductionist point of view.

[L]aws of nature do not supervene logically on the collection of particular physical facts.... [A] causal connection between two events is something over and above any regularity between the events. ... the very existence of such irreducible further facts raises deep questions about their metaphysical nature. Apart from conscious experience and perhaps indexicality, these constitute the only such further facts in which we have any reason to believe. It is not unnatural to speculate that these two nonsupervenient kinds, consciousness and causation, may have a close metaphysical relation. (Chalmers, 1996, 86)

I too propose a close metaphysical relation, but one in tune with Lewis's

reductionist account of laws rather than Chalmers' further-facts view. I will defend a "Humean Physicalism", so called because it is much like David Lewis's Humean Supervenience position but without the realism about possible worlds.<sup>5</sup> The Humean Physicalist I envision is particularly interested in Lewis's best-system account of scientific law. On this account, a principle counts as a law of nature just in case it is a theorem of the ideal systematic treatment of all occurrences; an ideal systematic treatment is one providing the best possible combination of *simplicity* of exposition with *strength* of descriptive content.

On Chalmers' view, contrary to Lewis's, laws and causes do not supervene on mere occurrences but are metaphysically distinct facts. For example, that *principle X is a law* is logically independent from claims that the occurrences unfailingly satisfy X, that X plays an important role in science, or that X is part of ideal science. Such "further fact" views of law seem intuitive to many. However, much work on Lewis's best-system approach to laws (Earman 1984, Halpin 1994, 1999 and 2003, Lewis 1994, Roberts 1999, Beebe 2000, Ward 2002, Loewer 2004, and Cohen and Callender, 2009) may be able to overcome the intuitions running counter to Humean physicalism. If so, I argue, there is room to press Chalmers' "close metaphysical relation" between law/causation and consciousness toward physicalism about mind. It is worth giving a hint about the argument. According to Lewis's account of laws, a "Hume-world"—a world defined to have occurrences exactly like the actual world's but to have neither laws nor causes—is not possible because a world with exactly these occurrences has the same best-system and so laws as the actual world. Yet a Hume-world would *seem* both conceivable and possible. Thus, the Hume-world problem for the Humean is much like the zombie problem for the physicalist: their coherent conceivability seems to show that there are "further facts" in the actual world beyond those of the Hume or zombie-world. With one stroke, I will argue, the perspectival best-system account resolves both problems: the supposed factual differences between the "worlds" in question amount merely to a difference of perspective (or interpretation) taken on the actual world.

Section II below describes a number of other "hard problems". It also begins an attempt to show how these problems involve conceptual novelty arising from perspective-dependent interpretation. Here Ronald Dworkin on "hard cases" of law and his interpretive conception of legal rights provide another important analogue. From a basis of constructive interpretation as integration, sections III and IV develop and motivate the perspectival best-system view as an interpretive account of probabilistic laws *and* mentalistic

---

<sup>5</sup> Just how possible worlds are taken less seriously will become apparent below. However, I do not have an exact characterization of Humean Physicalism. But it is an actualist view. A physicalist may want to say, e.g., that angels are possible in some sense but not that angels exist in some universe just as real as the actual universe.

theory, and so begin to spell out the close relation between the two. Section V provides a conclusion generalizing the results of III and IV.

## II

Chalmers' best-known concern is the "hard problem" of consciousness.

The hard problem of consciousness is the problem of experience. ...

[H]ow and why do physical processes give rise to experience? Why do not these processes take place "in the dark," without any

accompanying states of experience? This is the central mystery of consciousness. ... What makes the hard problem hard? Here the task

is not to explain behavioral and cognitive functions: even once one has an explanation of all the relevant functions in the vicinity of

consciousness – discrimination, integration, access, report, control – there may still remain a further question: why is the performance of

these functions accompanied by experience? (2002b, 247-8)

Thus, there would seem to be no way to "deduce" or fully explain consciousness from a physical/behavioral basis of biological structure and function. Jackson (1994), Chalmers (1996), and Chalmers and Jackson (2001) argue that *conceptual analysis* allows the deduction of many significant high-level facts from the totality of fundamental physical fact. For example, that water is H<sub>2</sub>O can be an *a priori* conclusion from premises about the totality of facts about atoms. Given that H<sub>2</sub>O is the dominant molecule in bathtubs, rain, the oceans, and that H<sub>2</sub>O can be shown to have the relevant physical properties to play the "water-role", its identification with water is *conceptually* entailed. Still, the problem of consciousness cannot be solved by such conceptual deduction: That there is experience at all associated with biological information processors requires a leap to a new and independent conceptual realm.

That is the plausible idea behind the problem. Still, there are a number of *other* "hard" problems in philosophy involving similar conceptual leaps. The earlier example involved nomic facts: facts about laws of nature and chance are known from data about occurrent facts but the nomic would appear to be a conceptual and explanatory leap beyond the occurrent. Moral and evaluative political truths apparently supervene on non-evaluative matters while being ontologically independent. Similarly, *law and jurisprudence* are grounded in societal facts (statues, judicial decisions, a constitution, and the intentions of their authors) but are often thought to transcend this basis.

Ronald Dworkin's account of societal laws and judicial interpretation provides an instructive analogy to scientific laws and scientific interpretation. Dworkin has long argued that there is more to law than positive law, that is, there are further legal facts than the those regarding statues, constitutions, legal decisions and the rest of settled common law. A "right answer" to a

difficult legal question (typically) comes not from closer reading of a statute or a historical understanding of judicial, legislative, or other authorial intent. Instead, Dworkin argues, legal decisions are properly based on “constructive interpretation” which requires seeing a social practice as “the best possible example of the form or genre to which it is taken to belong”. The legally best interpretation requires political coherence or “integrity”: The principle of integrity “instructs judges to identify legal rights and duties, so far as possible, on the assumption that they were created by a single author—the community personified.” (Dworkin 1986, p. 52)

One may say a bit more about the logic of interpretation. Dworkin’s own take on hard problems (of the law) is a good starting point. Dworkin concludes *Law’s Empire* with this.

Have I said what law is? The best reply is: up to a point. I have not devised an algorithm for the courtroom. No electronic magician could design from my arguments a computer program that would supply a verdict everyone would accept once the facts of the case and the text of all past statutes and judicial decisions were put at the computer’s disposal. But I have not drawn the conclusion many readers think sensible. I have not said that there is never one right way, only different ways, to decide a hard case. (Dworkin 1986, p. 412)

So, there may be right answers to hard questions of law, but understanding their correctness is not stereotypical conceptual analysis. Moreover, there is *some* reason in Dworkin to think that a “right answer” can be a matter of perspective.

Law as integrity, then, requires a judge to test his interpretation of any part of the great network of political structures and decision of his community by asking whether it could form part of a coherent theory justifying the network as a whole. No actual judge could compose anything approaching a full interpretation of all of his community’s law at once. That is why we are imagining a Herculean judge of superhuman talents and endless time. (Dworkin 1986, p. 245)....For every route that Hercules took from that general conception [of law as integrity] to a particular verdict, another lawyer or judge who began in the same conception would find a different route and end in a different place, as several of the judges in our sample cases did. He would end differently because he would take leave of Hercules, following his own lights, at some branching point sooner or later in the argument. (Dworkin, 1986, p. 412) [See also p. 239 where Dworkin is somewhat clearer about the correctness of other’s “lights” or what I call perspective. Dickson 2001 makes a similar point about perspective in Dworkin’s writing. Still, Dworkin is more explicit about the different standards for different *sorts* of interpretation. See the quote ending this section.]

For Dworkin, then, matters of law are *factual*; there can be a single right answer. But a right answer presupposes a *perspective* (one's "own lights" trading "off an interpretation's success on one type of standard against its failure on another", p. 239) and global thinking from that perspective. Judges are to give coherence (or integrity) to the law as a set of principles organizing the "community of principle". This derivation of principles is constructive and not a deduction or conceptual entailment of the Jackson-Chalmers variety.

Lewis's concept of *scientific* law is also an interpretive conception. His view takes laws to be a matter of the best way of understanding all happenings and again (as with Dworkin on societal law) in terms of principles best integrating this totality. Other accounts inspired by Lewis and John Earman's work may be even more to the point. The emerging account of laws is very context dependent in some versions including my own "perspectival best-system account" (Halpin 1994). The idea is still that a statement counts as a law of nature if it is a theorem of the best axiom system for all occurrences. But here "best" is relative to a perspective: plausibly biology utilizes a different perspective on systematization than physics. Most notably, evaluations of possible worlds are typically dependent on the perspective of the actual world. Because, this perspectival best-system view allows more than one "best" system assigned to a possible world, I one might call it simply the "systems view" of laws.

Dworkin too saw a connection between his idea of constructive interpretation for the social practice of law and other forms of interpretation, scientific included:

I offer this constructive account [of interpretation] as an analysis of creative interpretation only. But we should notice in passing how the constructive account might be elaborated to fit the other two contexts of interpretation [conversational and scientific] I mentioned, and thus show a deep connection among all forms of interpretation. Understanding another person's conversation requires using devices and presumptions, like the so-called principle of charity, that have the effect in normal circumstances of making of what he says the best performance of communication it can be. And the interpretation of data in science makes heavy use of standards of theory construction like simplicity and elegance and verifiability that reflect contestable and changing assumptions about paradigms of explanation, that is, about what features make one form of explanation superior to another. The constructive account of creative interpretation, therefore, could perhaps provide a more general account of interpretation in all its forms. We would then say that all interpretation strives to make an object the best it can be, as an instance of some assumed enterprise, and that interpretation takes different forms in

different contexts only because different enterprises engage different standards of value for success. Artistic interpretation differs from scientific interpretation, we would say, only because we judge success in works of art by standards different from those we use to judge explanations of physical phenomena. (1986, 53)

Dworkin calls the concepts that result from this process of making a subject matter S the best S it can be, "interpretive". As I understand it, this means that *interpretive concepts* (e.g., of legal rights) and their *rules of use* or *conceptual role* are generated in interpretation, i.e., are produced by giving a subject matter its best interpretation. Moreover, *truth* claims regarding interpretive concepts are grounded in facts about the best interpretation, e.g., that there is a constitutional right to privacy is a matter determined by the best interpretation for constitutional law. In the next section, I will briefly set out and motivate an interpretive conception of scientific law, *the systems view*.

### III

The standard best-system account as developed by David Lewis (1986, 1994) proposes that laws and objective single-case probabilities (chance) can be defined together: if the best systematic account of the physical world – of all occurrent fact – requires principles asserting the existence of chance, then these principles are laws and the chance claims are true. In this way, the best system is the truth maker for both laws and chance. This too is constructive interpretation. To use Dworkin's language, interpretive concepts of scientific law and probability make science the best systematizer it can be. Or to use Lewis's language reminiscent of Dworkin, laws are defined as the theorems of the best of the "integrated systems of truth" (Lewis, 1986, p. 123).

However, the best-system account of laws has seemed *intuitively* problematic. Is physical law just a matter of best system? Or, instead, is law a matter of further nomic facts in addition to the occurrent ones? Usually, intuitions favor an affirmative answer to the latter question so that matters of law are facts over and above the occurrent. Thus, intuition would seem to undermine the best-system account. One way to make out the difficulty is in terms of the Hume-world. Again, the Hume-world is a world exactly like the actual world with regard to occurrent fact but with no laws, no causality, and no other nomic connections. The opponent to a best-system account claims that this world is *possible* (if bizarre) and, of course, is distinct from the actual world. The standard best-system proponent must say (counter intuitively) that the Hume-world is impossible: Any world with exactly the same occurrences as those of the actual world will have the same best system, and so the same laws, as the actual.

There is, then, a clear analogy between Chalmers' arguments about



consciousness and the anti-best system argument just outlined. Both are committed to the possibility of an impoverished world – the zombie or Hume-world – worlds that are, respectively, physical and occurrence duplicates of the actual world but missing mental or nomic facts present in the actual. Accordingly, there are *further facts* (non-physical and non-occurrent facts) regarding the actual world. This result is deeply problematic for the Humean. I will argue that the perspectival systems view provides a solution to the best-system account's Hume-world problem and that this solution can be extended to undermine the zombie-world problem for a physicalist. A Hume-world, I will propose, is just the actual world interpreted from an odd perspective refusing to systematize. Normal perspectives systematize the occurrences with concepts appropriate for that perspective's level of complexity or abstraction. Then, for the case of the mental, just as the best-system proponent holds that law and chance claims may be true in virtue of their place in a best physics, claims about phenomenal and other mental properties are true in virtue of what ideal psychology says. Mentalistic concepts, I propose, are included in ideal psychology in part because they arise in a native "theory of mind" and so are very natural concepts, indeed *ineliminable from a normal human's perspective*. The zombie world, I will conclude, is just the actual world viewed from an especially odd perspective, one refusing to systematize mentalistically. I develop the perspectivalist systems view in 1. – 4. below. Its application to the problem of consciousness comes in section IV.

1. Recall that the Humean Physicalist presupposes basic facts involving mere occurrences; the world as postulated is, at basis, just one physical event after another. The Humean Physicalist's best-system account sees science as interpreting all occurrent fact in the best way possible, making of the data the best systematic whole it can be. Then, scientific laws are defined as logical consequences of this best axiom system.

As a first approximation, when a scientist theorizes she gives principles providing an account integrating the data in a reasonably elegant way, i.e., an account with as much descriptive force or *strength* as appropriately can be combined with *simplicity*. Theory, then, requires a tradeoff: the more one says about the data, the less simple the description tends to be. Scientific principles, on this view, are aimed toward a compromise providing *ideal summary*.

According to my systems version of the best-system account, best system is (unsurprisingly) very context dependent. In principle, there could be a number of systems counting as "best" but from different perspectives. Perhaps these different best systems will coexist the way contemporary biology and chemistry do (the models of biology are constrained by those of the chemistry), or on the other hand, even ideal best systems might conflict as contemporary quantum mechanics and general relativity do. In any case,

different sciences utilize different standards. More strikingly, we should expect ideal psychology or sociology to be “best” in a different way from a best physical system. In addition, perhaps, a Martian scientist would have very different standards of systematization. However, there is a special sort of context dependence emphasized by the systems view, a projectivism important for both theory of mind and an account of nomic necessity. The next subsection outlines this context dependence and its link to nomic necessity.

2. Think about a scientist working with a possible world/computer simulation of global warming; she starts with a model of the actual world up through the present time and extrapolates it into the future to make a possible world *W*. Because the model and the fully extrapolated *W* will be much simpler than the actual world, the best system for either can be quite different from that for the actual world. Thus, a proponent of the standard best-system account of laws may insist that *W* has laws different from those of this actual world. But such rethinking of *W*’s laws is not a part of scientific practice. Instead of axiomatizing anew, our climatologist naturally and for the *good practical reasons* of prediction and control, applies the actual world’s best system/laws (as best she can) to *produce W*. Nothing else would be informative about global climate change.

Two points deserve emphasis. First, the theorist’s job when modeling is to *project* laws (as they are understood at a time) onto a model under consideration. Thus, possible worlds (hypothetical situations, models) worthy of scientific consideration are extrapolations which use *actual* laws (as best we know them) to project a model into the future. Having extrapolated a model of the present through time by using these laws, one does not recalculate a new best system for the occurrences in *W* and then rethink how the model would evolve. Part of our perspective, then, determines how we understand possible worlds. If we have one world of special interest in our work – often the actual world – then we will project its laws (or supposed laws) onto possible worlds/models even when these worlds or models are clearly not actual.

Second, if we define the physically possible worlds as extrapolations – extrapolations using actual laws – then we have a notion of necessity of actual law based on the practicality of projecting our own best system onto models. I submit that this is all there is to physical or nomic necessity. “It is necessary that continued CO<sub>2</sub> production leads to significant global warming” reduces to the truth in all the relevant worlds, i.e., the extrapolated models. This is nomic necessity even an empiricist can love. (For details, see Halpin 1999.)

3. A best theory may include new concepts whose application may seem to involve facts beyond the occurrent. For example, quantum mechanics assigns probabilities to events. But probabilities are a matter of potentiality

and not mere occurrence. So, how might chance find a place in the Humean physicalist scheme? Following David Lewis, the Humean physicalist takes chance to be real in virtue of its inclusion in the best system: Insofar as chance claims *fit* actual statistics, they make the system better by neatly encapsulating the statistical data and, so, may count as true (provided their total system is the best).

Other non-occurrent concepts, interpretive concepts, may need to be a part of the best system. Quantum mechanics is often interpreted as including counterfactual as well as probabilistic claims. Other potentialities are included in physical theory, for examples force fields, potential energy, and electrical resistance. Like chance, I propose that these other non-occurrent concepts are real (or objective or vindicated) because they are a part of the best system. See (Halpin 2003) for details.

4. The best-system account has seemed intuitively problematic because, it is claimed, facts about laws exist over and above facts about occurrences. The apparent conceivability of a Hume-world, a world with all the occurrences of the actual but without laws, is just one illustration of this problem. As we have seen, the standard best-system account rules out a Hume-world: on that account, the totality of actual occurrences determines exactly one best system and so exactly one (presumably non-empty) set of laws.

However, the perspectival systems view makes sense of a Hume-world: The Hume-world is just the actual world viewed from a perspective that refuses to interpret it in any systematic way. From this perspective, the actual world has no laws. (A perfect intelligence might best represent the world as just a totality of occurrences; it has no need to simplify.) I propose to apply the same thinking to zombie worlds and the problem of consciousness. This will take some development.

#### IV

The normal human brain has the capacity to conceptualize its and others' brain states in mentalistic terms. According to much contemporary theory, we humans employ mentalistic concepts as part of a native and internalized conception of ourselves and other agents: a "theory of mind". There are many views on how to understand this theory of mind. Perhaps it is a human's own quasi-scientific theory to predict behavior, as the "Theory Theory" contends. At the other extreme, the simulation account, a human may understand another by using himself or herself as model. (The latter view is close enough to a theory of mind for current purposes. A scientific theory too can proceed by modeling. The systems account makes room for scientific theory that is not just axiomatization: Halpin 2003. See Nichols and Stich, 2003 and Goldman, 2006 for hybrid Theory-Theory/Simulation accounts.) In any case, I assume, a theory of mind akin to folk psychology

is operative in most humans.

One option for the systems account, the one I pursue here, is to suppose that the best mentalistic systematization will be based on this native theory of mind, much as physicalistic description is based on our native theory of objects. Then the Humean physicalist has a natural way to understand theory of mind: Its mentalistic concepts are interpretive ones that have a useful place in theory – much as chance helps systematize modern physics. Thus, beliefs, desires, selves, and states of awareness are real (or objective or vindicated) because of the place of the corresponding mental concepts in the best systematic account of our minds; “best” now with respect to the personal perspective that has evolved in humans. As such, this perspective is so natural that we cannot do without it. In what follows, I assume that ideal psychology is something of an *extension* of a native or internalized folk psychology. Call any such psychology, native, extended or idealized, a “theory of mind” or “TOM”.

Just *what* and *how* does an ideal TOM systematize, and, especially, how do its experiential concepts (i.e., the concepts of conscious awareness) help in this systematization? Because native psychology and its scientific extensions are far from ideal, the answers are not straightforward. Still, the genesis of TOM and its experiential concepts in natural selection is good reason to think that it is on to something. And one thing it may be on to, according to research surveyed below, is a systematizing ability importantly like that of chance. Some contemporary cognitive theory makes this very explicit: *conscious contents function to encapsulate or “summarize” relevant information within the cognitive system*. This is striking because the best-system account of law, chance, and other interpretive concepts of science is often stated with just that term, “summarize”. I develop this line of thought in the next two subsections.

1. On the matter of just *what* is systematized by theory of mind, contemporary psychology suggests that a native TOM is the brain’s way of representing its own and others information flow and processing *but from a personal point of view*. The evidence for this native system comes mostly from studies of particular brain deficits corresponding to disabled brain modules (or, quasi-modular neural networks) underlying TOM discriminations. Stich and Nichols (2003) describe the normal human as having mechanisms realizing a number of functions underlying a TOM:

- A perception detection mechanism...to detect what others see, hear, etc. (may be deficient in autistics)
- A discrepant belief detection mechanism (may be deficient in autistics)
- An agent detection mechanism (may be deficient in autistics)
- Percept detection self-monitoring mechanisms (may be deficient in schizophrenics)

- Propositional attitude self-monitoring mechanisms
- A possible-world imagination mechanism

The details, however, are not so important for present purposes. What is important is that there is evidence for an evolved system of neural mechanisms meant to interpret self and others from a personal-level.

How is a TOM valuable and so a candidate for a best system? First, native-TOM in a normal, mature human has significant evolutionary advantages. It takes other humans as agents with representations for world and selves. So, we humans get the native advantage of systematically representing self and others *as* selves: selves that represent or misrepresent, represent from a particular vantage point and state of knowledge and so can use the appearance/reality distinction to advantage. (It is good, for example, to know that you see something your competitor can not and even more important to know what one's cooperators know and plan.) Furthermore, the naturalness of theory of mind makes it valuable, indeed virtually essential, *for human thinking about other humans*. The worth of a system, according to the perspectival best-system account, is relative to its context. So, though some other intelligence might do without a TOM, normal humans cannot.

For current purposes, a TOM's experiential concepts are most important. What is their place and importance within a TOM? Stich and Nichols' "perception detection mechanism" provides hypothetical brain circuitry to determine their application. As well, folk psychology utilizes experiential concepts ("appearances are not always real", "when one dreams one is usually not aware of one's surroundings but is conscious only of one's imaginings", "seeing is believing", "pain is to be avoided", etc.) as does cognitive science when it attempts to explain experience or, at least, the "neural correlates" of experience. And, I am proposing, an ideal TOM (for humans) will continue to utilize experiential concepts to *interpret* brain information flow. This, I will argue, is analogous to the utility of nomic concepts, e.g., chance, in physical theory. For chance, we can see roughly how its postulation "fits" the facts (viz., conformance to actual statistics – chances neatly *encapsulate* the data<sup>6</sup>). How do the experiential concepts "fit" and help systematize? Any answer to this question will be premature given the early state of cognitive science's work on consciousness. Still, current cognitive science provides important hints about a TOM's experiential concepts and their value to mental systematization.

There are a number of cognitive theories associating or identifying neural and psychological states with consciousness. These include theories of attention, egocentric representations in mid-level brain processing (as

---

<sup>6</sup> Of course, the conformance is predictably imperfect for chance. Here I gloss over the difficulties with the notion of statistical fit.

opposed to further processed allocentric contents), contents of a virtual serial machine, second order representations, and information that is integrated and/or broadcast. Any or all of the processes on this short list could help set off a self-monitoring percept detector. All plausibly relate to a personal account of the self, an account produced by a TOM. Finally, all processes on this short list are plausibly very important ones to the brain's information processing. Insofar as these processes are important to a personal account, a TOM may well be ideal only insofar as it includes concepts sensitive to them. Perhaps, then, the TOM's experiential concepts encapsulate information about some or all of these aspects of brain information flow. (Note that this is *not* to say that a TOM's experiential concepts have extensions identical to brain states. Rather, these concepts help *interpret* brain information flow from the personal perspective so valuable for human interaction. Similarly, chance is not to be identified with frequencies or other occurrent facts; the chance concept helps interpret data from a perspective meant to compress the data by statistical methods.)

There is one especially broad and influential group of theories lumped together on this short list: *theories of brain information integration and broadcast*. Dennett's theory of consciousness falls within this very general category. On his view, brain contents with appropriate "clout" are conscious. Clout is something like "political power" within the brain: "When processes compete for ongoing control of the body, the one with the greatest clout dominates the scene until a process with even greater clout displaces it" (2005, 137). Dennett points out that the "hard question" for neuroscience is to see how this clout is worked out in the information processing of brains. The global workspace theory, GWT, is a contemporary to Dennett's view (Baars 1988) and now comes in many guises fleshing out potential answers to Dennett's hard question. On the GWT model and many closely related theories, contents are conscious when they are broadcast throughout the brain, i.e., when these contents are widely distributed to interact with many problem solving activities.

On some versions of this view, *integration* is explicitly included as a precursor to broadcast: the function of consciousness is to break the usual independence of various unconscious and habituated mechanisms, recruit their capacities for novel situations and problems, and thus integrate information from disparate and previously independent brain resources. Tononi (2004) provides a precise "information integration" theory of consciousness and argues that "the quantity of consciousness associated with a complex is determined by the amount of information that can be integrated among its elements, the quality of consciousness is determined by the informational relations that causally link it elements". Following Tononi, Koch has been a proponent of the "information integration" as (at

least) correlate to consciousness.

Consciousness is surmised to have substantially different functions from attention. These include summarizing all information that pertains to the current state of the organism and its environment and ensuring this **compact summary** is accessible to the planning areas of the brain, and also detecting anomalies and errors, decision making, language, inferring the internal state of other animals, setting long-term goals, making recursive models and rational thought. (Koch and Tsuchiya (2007), p. 17)

On this view, information integration – providing a *concise summary* – is a key to consciousness. I think it is clear, then, that summarizing could be a very important aspect of the brain's information flow, perhaps for all the reasons mentioned in this Koch and Tsuchiya quote. For example, contents that summarize the brain's perceptual information would be invaluable for broadcast to decision mechanisms. For present purposes, it should be stressed that the Tononi-Koch view provides the best model for taking experiential concepts to play a role analogous to chance in physical theory: Both chance and consciousness are seen as valuable for encapsulating information and thus may be essential for a best systematization.

Still, the science of consciousness is in its infancy. Though the Tononi and Koch account makes the analogy to chance especially vivid, there are other plausible accounts of consciousness (several mentioned above) and there may be far better accounts forthcoming. Ned Block's (2007) target article (and a number of his commentators) argue that the best "mesh" between psychology and neurophysiology entails conscious contents that are not accessible for integration and broadcast. I certainly do not mean to adjudicate. Rather, my main point should be relatively uncontroversial, that this best mesh – i.e., best systematization – of psychology with the rest of cognitive science will plausibly extend native theory of mind by utilizing experiential concepts. I then want to suggest that consciousness and other mental properties are real in virtue of this best system much as the proponent of the best-system account takes chance as real given its (putative) place in the ideal physical theory.

I have not yet argued that conscious states are *phenomenally* real. Even if ideal cognitive theory entails that information integration and broadcast plays an essential role in information processing, why is the word "conscious" apt? Can't information integration and broadcast – even as systematized by the ideal psychology – take place "in the dark" (to reuse Chalmers' metaphor)? In the next subsection, I will try to show how an ideal TOM may justify the claim that its experiential concepts are phenomenal and so, because they are also interpretive concepts, are the basis for consciousness.

2. Can the experiential/interpretive concepts of a TOM be *phenomenal*? Continuing to suppose that the idealized theory of mind is conceptually and theoretically an extension of current folk theory, one can see a number of reasons for the claim that an idealized TOM provides a basis for phenomenality, i.e., the reality of qualia. *First*, its experiential/interpretive concepts are *subjective* – this is no surprise as folk psychology can be subjective. Think about such judgments as “the apple looks red to me” (or “I see it as red”, “it has a red look”, etc.) A TOM plausibly implies that if the apple looks red to me, then it probably is red, but that sometimes such appearances are deceiving. Insofar as claims like these are entailed by an ideal TOM, the *self and subjectivity* are real in virtue of their places in the best system. This constructive interpretation is a “package deal” akin to Lewis’s inter-definition of “law” and “chance”. *Second*, the content of the concepts of a TOM can be of the appearances. The “look” of the apple is a quality of experience: a *quale*. Again, these are real in virtue of their place in ideal personal theory. We should expect an ideal TOM to have much to say about sensing as well as believing. Furthermore, this account may help explain the unified character of qualitative experience (insofar as consciousness is information *integration*) and its privacy (as experiential discriminations derive primarily from the perspective of self-monitoring mechanisms).

*Third*, an idealized TOM need not identify subjective, qualitative states (or their tokens) with any particular brain or physical state (or token). Analogously, an idealized physical theory invoking chance need not identify probability for event type x with any particular frequency or pattern of x-occurrence. Insofar as the qualia postulation by a TOM is likewise an interpretive conception, the metaphysical status of the phenomenal may mirror the autonomy of chance from actual relative frequencies. The properties derived through constructive interpretation supervene on the physical occurrences but need not be related in any specific way to their supervenience base. Figuratively, the interpretation of a candidate best system provides an overlay re-conceptualizing the supervenience base of occurrent, physical fact. There is no need for an identification of the new properties with the more basic properties or functions among these. Thus, *fourth*, interpretive concepts may arise without structural, historical, or functional definition. Their implicit definition in best systematization arguably means that chance and qualia are conceptually intrinsic (independent of explicit definition). This is in accordance with intuition and standard views of both chance and qualia.

It is worth recapitulating the chance-consciousness analogy. An interpretive scheme for mere occurrences, an ideal physical theory, is the truthmaker for statements about chance and so for a kind of potentiality. Likewise, the personal perspective’s ideal theory of mind, systematizing



brain information flow, is the truthmaker for statements about consciousness. Still, I suspect the systems account of consciousness to be even more controversial than that of chance. This is no surprise on the present account. Because TOM-discriminations are mostly automatic and largely unconscious, phenomenality is mysterious in a way that systematizing in terms of chance is not: Truths about chance are seen to be confirmable by experimental data while truths about consciousness are immediately intuited through native self-monitoring systems. On the present view, then, consciousness emerges on interpretation just as chance does; the salient difference is that the systematizing involved in our native theory of mind is mostly unconscious.

Autistics, on a standard account, have an incomplete theory of mind due to defects in their TOM mechanisms. Thus, they sometimes cannot see humans as full-fledged persons. From *their* perspective, all of us are something like zombies. In any case, I suggest that the zombie-world is just our world but from a perspective allowing *no* theory of mind conceptualization. Just like a Hume-world, a zombie-world is our world but from a perspective in which normal systematization is suspended.

## V

Finally, there is reason to believe that the systems view makes good sense of special science law more generally. Callender, C. and Cohen, J., (2010) apply their "better best-system" account (roughly what I have called *the systems view* or *the perspectival best-system account*) and show in detail how special sciences are accommodated.<sup>7</sup> On their view as on my own, a special science may have a best systematization. Still, "best" is relativized to the perspective determining both the basic concepts and the standards for simplicity-strength tradeoff given these concepts. Biological science may be more interested in trading off simplicity for greater strength of description within its domain; after all, there is much to describe given the many species existing in our small part of the universe. But the biological theory is, I propose, still a systematization of occurrent physical fact. Similarly, recapitulating section IV, a theory of mind fleshing out folk psychology also systematizes. It takes a native, personal perspective with non-physical, non-spatial conceptualization. Still, so long as the basis of this theory of mind is to brain information-flow occurrences, then the Humean physicalist sees it as a candidate best system.<sup>8</sup> (Just how a personal-perspective account is part of a best-systematization of brain information-flow is a matter of much cognitive science work going on since this paper

---

<sup>7</sup> They, however, require supervenience on the kinds of ultimate physics. I propose supervenience on physical occurrence.

<sup>8</sup> That the brains occurrences to be systematized are informational presupposes an interpretation on the brain events. So understood, I would argue that there is no problem here for the Humean physicalist.

was originally written.)

Also, much like cases surveyed earlier, the systems account is happy to project the personal perspective where it proves useful, e.g., onto animals, even some failing to have a theory of mind of their own.<sup>9</sup> Of course, just which organisms count as conscious depends on what is in fact the best TOM. (It is clearly not best to apply mental concepts non-discriminately.)

Systematization for different sciences is, thus, perspective-relative. It depends on efficient integration of the data given the concepts and standards of that science. As we have seen, Dworkin, Lewis, and Tononi explicitly appeal to this integration. The systems account is clear on why. Whether providing a constructive interpretation justifying legal propositions (Dworkin), systematizing the occurrent physical facts (Lewis), or integrating brain information flow (Tononi/Koch), a good theory projects useful, perspective-dependent concepts, together with standards of coherent systematization, onto a subject matter. Useful and coherent systematization is in the eye of the beholder or, rather, is relative to the needs of the consumer. Thus, different sciences utilize different information compression schemes for their particular purposes. The moral I draw is that special science concepts and their true ascription is a matter of perspective-dependent systematization (or interpretation) rather than of “further facts” beyond the physical.

---

<sup>9</sup> Much as the systems account endorses projecting the actual world’s best physical system onto physically possible worlds, a theory of mind is to be projected onto possible scenarios in which the creatures or machines in question have no theory of mind of their own.

## REFERENCES

- Beebe, H. (2000) "The Non-governing Conception of Laws of Nature", *Philosophy and Phenomenological Research*, 2000: 571-94.
- Block, N. (2007) "Consciousness, Accessibility and the Mesh between Psychology and Neuroscience," *Behavioral and Brain Sciences* 30: 481-548
- Cartwright, Nancy (1999): *The Dappled World. A Study of the Boundaries of Science*. Cambridge: Cambridge University Press.
- Chalmers, D.J. (1996). *The Conscious Mind*. New York: Oxford University Press.  
(ed.)
- Chalmers, D.J. (2002a) *Philosophy of Mind: Classical and Contemporary Readings* (Oxford).
- Chalmers, D.J. (2002b) "Consciousness and its Place in Nature" in Chalmers (2002a)
- Chalmers and Jackson (2001) "Conceptual Analysis and Reductive Explanation". *The Philosophical Review*, 110.
- Callender, C. and Cohen, J. (2010) "Special Sciences, Conspiracy and the Better Best System Account of Laws", *Erkenntnis* 73: 427-447.
- Cohen, J. and Callender, C. (2009) *Philosophical Studies* 145(1): 1-34.
- Davidson, D. (1987) 'Knowing One's Own Mind', in *Proceedings and Addresses of the American Philosophical Association* 61: 441-58.
- Dennett, Daniel (2005) *Sweet Dreams: Philosophical Obstacles to a Science of Consciousness* The MIT Press.
- Dickson, Julie (2001) "Interpretation and Coherence in Legal Reasoning", *The Stanford Encyclopedia of Philosophy* (Fall 2001 Edition), Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/fall2001/entries/legal-reas-interpret/>>.
- Dworkin, R (2011) *Justice for Hedgehogs*, Belknap Press of Harvard University Press, Cambridge, Massachusetts.

- Dworkin, R (1986) *Law's Empire*, Harvard University Press, Cambridge, Massachusetts.
- Earman, J. (1984), "Laws of Nature: The Empiricist Challenge", in Bogdan (ed.), D.M. Armstrong, Reidel, Dordrecht, Holland.
- Halpin, J.F. (forthcoming) "Fine-Tuning Arguments and the Concept of Law"
- Halpin, J.F. (2003) "Scientific Law: A Perspectival Account" *Erkenntnis*; 58(2): 137-168.
- Halpin, J.F. (1999) "Nomic Necessity and Empiricism", *NOUS* Vol. 33: 630-643.
- Halpin, J.F. (1994) "Legitimizing chance: The best-system approach to probabilistic laws in physical theory", *Australasian Journal of Philosophy*, 72:3, 317-338
- Jackson, F. (1994), "Finding the Mind in the Natural World", reprinted in Chalmers (2002a).
- Koch, C. and Tsuchiya, N. (2008) "Attention and consciousness: two distinct brain processes" *TRENDS in Cognitive Sciences*, Vol.11 No.1.
- Kripke, S. (1980) *Naming and Necessity*, Harvard University Press, Cambridge, Massachusetts.
- Loewer, Barry (2004) "David Lewis's Humean theory of objective chance," *Philosophy of Science* 71: 1115-25.
- Nichols, S and Stich, S. (2006) *Mindreading*, Clarendon Press, Oxford.
- Lewis, D. (1973) *Counterfactuals*, Harvard University Press, Cambridge, Massachusetts.
- Lewis, D. (1986) *Philosophical Papers*, v. II, Oxford University Press, Oxford.
- Lewis, D. (1994) "Humean Supervenience Debugged", *Mind*, Vol. 103.412: 473-490.
- Roberts, J. (1999) "Laws of Nature' as an Indexical Term", *Philosophy of Science*, 66 (3 Supplement): S502-S511.

Tononi, G. (2004) "An information integration theory of consciousness", *BMC Neuroscience* 5:42.

Tononi, G. and Koch, C. (2008) "The Neural Correlates of Consciousness: An Update" *Ann. N.Y. Acad. Sci.* 1124: 239–261.

Ward, B. (2002) "Humeanism without Humean Supervenience: A Projectivist Account of Laws and Possibilities", *Philosophical Studies*, 107, 191-218.