

1

Conventionalism: Epistemology Made Easy, Ontology Made Paradoxical

We manage, it seems, to learn much about the kinds and stuffs and phenomena which surround us in nature. Through attentive inspection of individual members of a given kind (or individual samples of a given stuff, or individual instances of a given phenomenon), we manage to identify properties which all members of the kind are bound to possess, so long as they exist at all. Among these are often properties, or combinations of properties, which members of no other kind can possess. But exactly *how* do we manage to identify essential natures, distinctive of nature's various kinds, stuffs, and phenomena? From what premises do we infer such conclusions? The only developed answer to this question currently on offer leads to unsettling conclusions about the ontological status of essential properties. Or to speak more precisely, it leads to unsettling conclusions not about the properties themselves that we determine to be essential to nature's kinds, but about the ontological status of their *being* essential, of their essentialness. In this chapter I will argue that these conclusions are not just unsettling but unbelievable. In the next chapter I will offer an alternative answer to the epistemological question.

Just what evidence apprises us that chromium necessarily has an atomic number of 24, that quartz crystals by nature have their molecules arranged in a certain sort of lattice, that lightning is essentially an electrical phenomenon? Not just that inspected samples uniformly present the property in question. For we manage to draw distinctions between properties which samples of a stuff or members of a kind uniformly possess, and properties which they possess by their very nature. We determine that all samples of chromium come originally from Zimbabwe or South Africa or Siberia,¹ but do not judge that coming-from-Zimbabwe (-or-South-Africa-or-Siberia) belongs to chromium's very nature—that the samples *had* to come from Zimbabwe or South Africa or Siberia. We may learn that diamonds are all marketed by a monopoly enterprise, but do not infer that they are by nature marketed in this way; we distinguish between their *being marketed by a monopoly* and their *being composed of carbon*. To put it differently, we somehow learn that counterfactuals beginning "If chromium had been present in the United States, . . ." may have completions that make them true and important from the standpoint of geology or economics or politics, whereas counterfactuals beginning "If chromium had had atomic number 79, . . ." are empty and uninformative—true only vacuously. But how?

Kripke (1972) famously argued that we learn from experience that gold essentially has atomic number 79, water necessarily has molecular structure H_2O , and that (degree of) heat is by nature (degree of) mean kinetic energy. These properties are, as science informs us, explanatorily rich—they explain other properties that gold and water have with equal uniformity, or enable us to predict uniform connections between (degree of) heat and pressure. This encourages the thought that explanatory richness is the extra

premise. If samples of kind K uniformly bear property p , and p is in this way explanatorily rich, perhaps it follows that p is an essential property of kind K . But is it not also an essential property of gold that gold has a melting point of 1073°C , or that gold resists corrosion by all acids and acidic compounds except aqua regia? Perhaps explanatory richness is not a necessary condition for a property's being essential. For that matter, it does explain a good deal about diamonds that diamonds are marketed by a monopoly enterprise. It explains why they are expensive, and perhaps thereby explains why they are given as tokens of important occasions or deep feelings, and so forth. "Explanatory richness" of at least some sorts may not be sufficient—even when added to uniform occurrence—to ensure essentialness.

1.1 Do We Know "Template" Truths about Essential Natures?

So what *is* the extra premise that, when added to the uniform occurrence of p among inspected members of K , permits us to infer that p is an essential property of K s? The only answer to this question that is now widely defended is that we combine the uniformity we empirically discover among members of K with something we somehow know about the *kind* of kind to which K belongs (McGinn 1981, pp. 157–158; Sidelle 1989; cf. Jackson 1998 on our knowledge of C-extensions). Thus it is said that we know, concerning chemical compounds such as water, that whatever the molecular structure that samples of that kind prove uniformly to possess, it is a molecular structure that samples of that kind *essentially* or *necessarily* possess.² It is said that we know, concerning physical elements, that if samples of physical element K prove uniformly to have atomic number x ,

physical element K has atomic number x by nature. Thus do we know that the only nonempty counterfactuals concerning K must depict K as having—or at least must be consistent with K 's having—just that atomic number. It might be said we know, concerning the substances which the mineralogist studies, that whatever the molecular arrangement that the mineralogist determines samples of such a substance uniformly to have, it is an arrangement that instances of that substance are bound to have—as quartz is bound to have a particular lattice arrangement.

But how do we manage to know these “template” truths concerning the kinds of nature's kinds? Do we learn them from experience? The thought here would have to be that we perform a metainduction. We first infer from induction over samples of gold that gold has atomic number 79 essentially; from induction over samples of chromium that chromium has atomic number 24 essentially; and at length do a metainduction over physical elements in general—inferring that each of them is characterized essentially by a particular atomic number. But this thought obviously cannot be defended, at least not in just this form. For we cannot on the present way of thinking even arrive at the premise that gold is characterized essentially by atomic number 79 unless we already know that conclusion to which the metainduction is to lead—that physical elements select an atomic number not just uniformly but by nature.

Then how might we be said to know of these template truths? One answer might be that we exercise a direct intellectual insight, not mediated by experience, into the natures of the higher-order kinds—for example, “physical elements,” “chemical compounds,” “mineral substances”—into which nature's specific kinds fall. But that answer seems fanciful, of course.

Or might we learn of such template truths by armchair reflection on our own classificatory practices? Perhaps it is just a convention of ours to individuate physical elements by atomic number, chemical compounds by molecular composition, and mineral substances by (chemical composition and) molecular arrangement (Sidelle 1989). Perhaps, that is, it is our convention *not* to judge or say that the same chemical kind is present in two envisioned scenarios—two actual contexts, or two counterfactual contexts, or a mix—unless the kind envisioned in both is envisioned as having a single molecular composition. If we do have conventions of individuation such as this, it seems plausible that upon armchair reflection we would sense *that* we have them. We would find ourselves being drawn to deny that a look-alike of water, envisioned from the armchair as existing in some scenario, were *the same* chemical stuff as water as soon as we realized we were envisioning this look-alike as having a molecular composition other than H_2O —for example, the molecular composition abbreviated as “XYZ.”

1.2 Conventionalism, and Essentialness as Mind-bestowed

But if this is how we arrive at our judgments that certain properties characterize nature’s kinds not just uniformly but essentially, conclusions follow that are at least disturbing. Are these judgments truly warranted? The extra information we are now pictured as adding, to the empirical finding that (say) gold uniformly displays atomic number 79, is that we will not *call* a physical element “gold again,” in speaking of an envisioned counterfactual scenario, unless we manage to envision that element as having the same atomic number as we have empirically identified in actual

samples of gold. Is this enough to warrant the conclusion that a physical element cannot or could not *be* gold, unless it had just that atomic number? Does the fact that we would not *call* something “gold” warrant the conclusion that that something could not *be* gold? Well, perhaps the conventions governing what we will *call* “gold”—our conventions for individuating, our practices of classification—are the way they are for a reason. Perhaps they have somehow been shaped by the way the world is. But to call something a “convention” is to suggest that we had latitude in adopting it—that we could have proceeded differently. Now it is true that that suggestion is avoided if we speak instead of our *practices* of classification. But still there is nothing in the view we are examining that suggests that our practices are shaped by empirical contact with the world—and hence nothing, barring the answer scotched above as fanciful, that suggests that they are shaped by the way the world is.

So if our judgments of essentialness are truly to be warranted, on the view we are examining, our conventions for calling something “the same kind again” must be seen, not as *evidence for* its being the same kind again, but as *constitutive of* its being the same kind again (Sidelle 1989, pp. 49, 65, 67). That members of a given kind must cling to certain properties through thick and thin, in all actual phases of their careers and in all counterfactual scenarios, must not be something *indicated* or *suggested* by our conventions’ being such as they are, but something that obtains *in virtue of* our conventions’ being such as they are. The essentialness of essential properties is essentialness relative to us, relative to our conventions or practices. The essential status of essential properties is mind-dependent.

What is disturbing about this result is the way it intersects with the thought that the essential properties of members of

a given kind are properties which those members must retain, so long as they exist at all. If the lattice arrangement essential to quartz crystals is removed, then where a moment ago there was a quartz crystal, there will be a quartz crystal no longer—the quartz crystal will have been destroyed. If the property of containing-79-protons-in-the-nucleus disappears, where a moment ago there was a gold atom, that gold atom will have ceased to exist. But in virtue of what are these occurrences *destructions*—ceasings-to-exist—instead of mere *alterations* in something that continues? This is the same question as: in virtue of what are these *essential* properties of quartz crystals and gold atoms, and not just properties that quartz crystals and gold atoms have so far proven to have? And the answer on the present view will be: in virtue of our conventions' being such as they are (Sidelle 1998, pp. 440–441). Independently of us, there will be in the world only a play of properties, one property giving place to another and that property to another in turn. That some switches of properties amount to ceasings-to-exist, that others amount to comings-into-existence, whereas yet others amount to mere alterations, is the case only relative to us and our conventions. In other words, that the existences of the world's objects begin where they do, and end where they do, will not be independent of us and our conventions. Beginnings and endings of existence, for the world's objects, will obtain only relative to us.

Should this result be articulated in antirealist fashion, as the claim that we by our conventions actually *construct* the existences of the world's objects? Proponents of the conventionalist account of essentialness in fact divide on this question. Alan Sidelle, a prime exponent of conventionalism, provides an austere interpretation that avoids anti-realism; many other exponents elect strongly antirealist

formulations. My position is that either style of articulation is disturbing in its own way—indeed, if the arguments of the next section are right, conventionalism on either articulation is simply not believable.

On Sidelle's version of conventionalism, all that there is in the world, independently of us, is "stuff" (or, as a commentary on Sidelle calls it, "world-stuff").³ World-stuff is by no means undifferentiated: it bears all manner of different properties, and throughout it particular properties routinely get replaced by other properties. But there are no objects in the world as it is independently of us. For objects are (or would be) entities that get destroyed when certain properties are replaced, and merely alter when certain other properties get replaced—objects have certain properties essentially, and others merely contingently (Sidelle 1998, p. 441; 1989, p. 55n.). And there are, in the world as it exists independently of us, no modally qualified states of affairs. Apart from world-stuff—apart from the world as it exists independently of us—there is only us. That is, there are our conventions of individuation, and (presumably) the utterances and thoughts that implement these conventions. Our making these utterances and having these thoughts create in us the impression that there are in the world objects, having certain properties essentially, but this impression is strictly false. It must be added that Sidelle's writings are tight-lipped about just what *our* existence involves—it may, for all the texts show, amount to no more than the occurrence of a series of such utterances and thoughts.

Most philosophers of generally conventionalist sympathies elect a richer picture of the world. There do exist in the world objects, on the richer sort of picture, and the existences of the world's objects have beginnings and endings. But they have these only relative to our conventions for indi-

viduation, our practices of classification. So it is in a sense true that we (by our conventions) construct, shape, the existences of the world's objects. This is of course "construction" in a transposed sense. It does not require the use of hammers and saws, and we do not do it in the sweat of our brows. We do it merely by thinking and talking as we do. And if this sounds mysterious—how, by just thinking, can we make objects arise and last for determinate periods and then cease to exist?—the answer is that the objects to which we do this are as insubstantial as our own constructing activities. They have only the shadow reality of a mental (or a linguistic) projection. But *being* just that—having no existence save existence-relative-to-our-thought-and-talk—they really are entities whose existences we delimit just by thinking and talking. There are in the world no "ready-made objects" (Putnam 1982; cf. Putnam 1981, pp. 53–54). Rather the world of objects is "a kind of play," a series of stories, of which we are the authors; we do ourselves appear in the stories, but nevertheless "the authors *in* the stories are the *real* authors" (Putnam 1977, p. 496).

1.3 How Conventionalism about Essentialness Yields Paradoxes

Are these two alternative conventionalist pictures of the world not just unsettling—or exciting, depending on one's point of view—but outright untenable? That is what I now will argue. I will begin with a paradox that confronts at least many, probably most, conventionalists who elect the antirealist picture. I will then present two parallel paradoxes, one of which confronts the rest of the conventionalists who elect the antirealist picture, the other of which confronts conventionalists who elect the austere realist picture of Sidelle.

Most philosophers nowadays subscribe to a materialist view of our minds: human mental events are by nature events befalling human brains. It is fair to infer that many, and probably most, conventionalists are committed to this general view. There are of course importantly different versions of materialism. Some hold that our mental events are brain events neurochemically specified, others that they are brain events functionally specified, yet others that they are brain events teleofunctionally specified (Millikan 1984; Elder 2001b). But all materialists—including all conventionalists who are materialists—are committed to the position that the existence in the world of human brains is logically prior to the occurrence in the world of human mental events. Human mental events are by nature events that happen in or to human brains; unless and until there are human brains in the world, there can occur no human mental events.

Yet human brains seem *par excellence* to be entities that can survive some alterations and cannot survive others; they seem to have essential properties, properties they must retain if they are to go on existing at all. Just what are those essential properties—to what natural kind do human brains belong? In chapter 7 I will present reasons for thinking that human brains all by themselves amount to a particular natural kind. But even there I will defend only general remarks about the *kinds* of properties that characterize them essentially. Specific answers on the properties essential to human brains is a question for empirical science, I will argue. Still it is safe to say that human brains must retain certain properties of structure and organization if they are to go on existing at all. A human brain cannot survive being compressed to the size of a sugar cube; it will likewise be destroyed if a bolt of lightning vaporizes it and disperses its

component molecules. If human brains exist in the world at all, there exist in the world entities that essentially have a certain structure and organization.

Suppose then that some human brain undergoes a change that removes some of these properties of structure. In virtue of what is this change a destruction—an end of an existence—rather than merely an alteration in something that continues to exist? Conventionalists—at least, conventionalists who believe there are in the world objects—must answer: in virtue of our having the conventions of individuation that we have. But our having our conventions is a matter of our thinking and talking in certain ways. It is a matter of our undergoing certain mental events. So the occurrence in the world of at least some human mental events is logically prior to the existence in the world of human brains. For it is in virtue of our conventions that there are in the world entities having essentially the properties of structure that human brains have essentially.

Thus conventionalists who are materialists must say: the existence in the world of human brains is logically prior to the occurrence in the world of human mental events, and the occurrence in the world of human mental events is logically prior to the existence in the world of human brains. This is a paradox.⁴ And by “paradox” I do not mean a pleasant puzzle about which to spin articles. It is a paradox in the original sense—it is *para doxa*, beyond belief.

Can conventionalists who believe that there are in the world objects—conventionalists who eschew Sidelle’s austere realist picture—save themselves by embracing dualism? But even dualists must claim that there are certain changes that human minds can survive, and others that they cannot. A human mind can pass from entertaining one thought to entertaining another without ceasing to exist. But

a human mind cannot acquire an atomic number of 79 or a valence of +3. If there suddenly arises, where a moment ago there was found a human mind, an entity having atomic number 79, then there a human mind has ceased to exist—surely even a dualist must agree with this. But in virtue of what is a change in the thought entertained merely an alteration in a human mind—merely a switch in properties accidental to a human mind—while drastic alterations like the one just considered amount to the ending of a human mind's existence? Conventionalists who are dualists must answer: in virtue of our having the conventions of individuation that we have. So the occurrence in the world of human mental events is again logically prior to the existence of human minds. But isn't the occurrence of human mental events logically posterior to the existence of minds that can undergo them? I shall assume that any dualist must answer Yes—that any dualist must deny that mental events can occur logically prior to, and independently of, the existence in the world of minds.

But then *any* conventionalist who believes that there are in the world objects—any conventionalist electing the anti-realist picture—is caught in a paradox.

What then of Sidelle's austere realist picture of the world? Here there are no objects, no courses of existence, no distinctions between mere alterations and outright destructions (or creations). There is only world-stuff, on the one hand, and on the other hand us and our conventions of individuation.

But let us ask: why are the conventions of ours, in virtue of which some properties of the objects which we believe in are essential, and other properties merely accidental, called "conventions of *individuation*"? Because there is a close connection between our individuating as we do and our affirm-

ing the modal judgements that we affirm. Thus far we have observed this connection only at the level of kinds. We have noted that, for Sidelle, our conventions forbid us to classify any substance envisioned in some imagined world as being “the same chemical stuff” as the water with which we are familiar unless we are prepared to think of that substance as sharing the same microstructure that familiar water has. We sense that this is our convention, and articulate the awareness by asserting that water takes microstructure H_2O with it through all possible worlds—that water essentially has that microstructure—since we incautiously suppose that *there is* in the world water, and other stuffs such as water, and thereby are required to suppose that there are in the world necessities.

But the connection between individuation and modal commitments obtains at the level of individual objects and samples as well. In order to judge that there exist, at the same time, *two* individual *Ks*, we must believe that there exists at that time a *K* having some property *p*, and a *K* having some property *p'*, such that no one *K* can simultaneously have *p* and *p'*. The clearest example of such thinking involves spatial locations. We typically are prepared to judge that there *now* exist in the world two objects of a space-taking sort *O* if and only if we suppose that there now exists an *O* having spatial location *s*, and an *O* having location *s'*, such that no one *O* can at a time have both *s* and *s'*. And it is in general easy to suppose this: with rare and strange exceptions, we suppose that extended objects of any kind necessarily cannot simultaneously occupy two discontinuous spatial regions. Almost as familiar are examples of analogous thinking involving temporal locations. Might there *here* exist, over the course of the world's history, two distinct individual *Ts*—two entities of a kind that enjoys

temporally extended existence? With rare and strange exceptions, it is a sufficient condition for our judging this that we suppose there here exists, over history, a *T* having a career that spans certain times, and a *T* having a career that spans other times, such that no one *T* can exist across both spans. And supposing this is at least often necessary for our judging there here to exist over history two *T*s. To suppose this is typically easy: almost without exception, we suppose that no time-taking object can exist across temporally discontinuous spans of time. In the cases where we need *not* be persuaded of such temporal discontinuity, to judge that there here exist over history two distinct *T*s, that will be because we suppose the *T* existing here at the later time had some one property, and the *T* existing here at the earlier time had some other property, such that no *T* can over its lifespan have both.

But the point is wholly general, and applies even to entities not located in space and time. We treat it as a necessary and sufficient condition, for there to exist in the world two (or more) *X*s, that there exist in the world an *X* having some property *p*, and an *X* having some property *p'* (etc.), such that no one *X* *can* have both *p* and *p'*. Joint possession of *p* and *p'* must be impossible for *X*s—it must be something that *X*s by nature cannot do, something incompatible with what *X*s essentially are like. The occurrence of a plurality of individuals of the same type, our conventions of individuation say, involves the obtaining of incompatibilities-with-some-essential-nature.

But the worldview Sidelle offers us holds that it is by virtue of our existing, and having the practices of individuation that we do, that there appear to be in the world any necessities and any essences—and that appearance is, moreover, deceptive. It seems fair to ask: in virtue of what are we a “we”—a plurality of minds—and in virtue of what are our

conventions of individuation plural? Is there nonconventional or preconventional individuation in the world? If so, our conventions of individuation are not the sole ground of “necessities,” and necessities are not mere appearances. If not, the obtaining in the world of our conventions of individuation is logically prior to the existence of *us* as a plurality—and for that matter is logically prior to the conventions’ being *conventions*, plural. Yet surely it must *also* be true that our *existing* in the world is logically prior to our having any particular conventions.

I conclude that even Sidelle’s version of the conventionalist position is *para doxa*—beyond belief.

1.4 Lewis-style Conventionalism

Before closing this section, and while we are still on the topic of plurality and individuation, I will comment briefly on a variant of conventionalism that holds that there are different correct answers, depending on the conversational context, as to which properties, or which *sorts* of properties, are essential to a given stuff or kind or individual. This is David Lewis’s “counterpart theory” about essential properties (Lewis 1986b, pp. 248–263). Lewis holds, as is well known, that there are countless real worlds in addition to the actual world. Hence any individual object in the actual world is significantly similar, in one respect or several, to countless nonactual objects across this range of worlds. The samples of any actual stuff, the members of any actual kind, will likewise all be similar to countless sets of otherworldly samples or otherworldly kind-mates. So we have in principle a great deal of latitude as to *which* otherworldly objects we will treat as *counterparts* to a given individual object or kind or stuff—as truthmakers for statements about ways the given object or kind or stuff could possibly be, even though

it is not actually that way. Yet the interests and presuppositions we bring with us, to any given conversational context, will limit what can count there as counterparts. “Two things may be counterparts in one context, but not in another; or it may be indeterminate whether two things are counterparts” (1986b, p. 254). The right thing to say, about which of the properties of an individual or kind or stuff are essential to it, will then be just as shifting, as subject to indeterminacy, and as context sensitive as is the extension of the counterpart relation. If a property possessed by a given individual or kind is missing in some of the contextually relevant counterparts, that property is accidental to the individual or kind; if the property is possessed by all relevant counterparts, that property is essential. In different contexts, different answers will be correct as to which properties are essential and which are accidental.

It takes a moment to understand just what this view of Lewis’s is a view *about*. Is it a view about what it is for properties to be essential to an individual or a kind or a stuff? On the traditional conception, the properties essential to an individual are properties it is by nature incapable of losing; those essential to a kind or a stuff are properties that any member of that kind, any sample of that stuff, is by nature incapable of lacking. Can it happen that an individual or kind or stuff should lack any of the properties which it is by nature incapable of lacking? No, that is a contradiction in terms. Can it happen that a given individual or kind should be or become capable of lacking properties that it by nature is incapable of lacking? No, that too is a contradiction in terms.

So any theory that says that the properties essential to a given individual or kind differ, relative to different contexts, is not a theory about what it is for properties to be *essential* at all. I infer that Lewis’s view is *not* about what it is for

properties to be essential, and that Lewis does not believe that, strictly, objects have any properties essentially. No, the real topic of Lewis's view must be what I suggested at the outset: it is a theory about correctly *saying* which properties are and are not essential to a given individual or kind or stuff. Our *saying* that these and those properties are essential to this or that individual or kind has to be the root phenomenon, on Lewis's view. A given property's *being* essential to this or that individual or kind has to be merely the flickering shadow, the inconstant projection, of the sayings that are required of us, in the conversational context, by our interests and customs and conventions. The latter render certain attributions of essential status *correct*. But no such attribution is ever literally *true*.

Lewis's view then is a variant of conventionalism, a projectivist view about essential status. Should we think of it as reflecting an austere ontology, like Sidelle's, on which there are only our sayings and a neutral world-stuff? Or should we think of it as an antirealist view, on which there are objects in the world, but projected objects, objects whose careers we construct? Lewis's texts comport better with the latter interpretation, but there are difficulties, as we have seen, with either alternative. Yet there are additional difficulties for Lewis, I suggest, connected with the very claim that there is a *plurality* of conversational contexts. The contexts evidently are plural independently, and prior to, the being-correct of *any* attributions of essential status. It is fair then to ask: in virtue of what are the contexts *contexts*, in the plural; what constitutes their distinctness from one another? If conversational contexts were distinct from one another in virtue of bare haecceities alone, it could not be explained how we learn which conventions and practices apply in *this* context, and which others apply in *that* context. The contexts

must then be qualitatively different from one another. The distinctness of context C_1 from context C_2 must rest on the fact that property p_1 is somehow involved in C_1 , property p_2 is involved in place of p_1 in C_2 , and p_1 and p_2 exclude one another—no single context can feature p_1 and p_2 in just the same role. But the only way of spelling out “in place of” or “in the same role as” is to identify property bearers common to C_1 and C_2 , property bearers that can have p_1 and can have p_2 but cannot, while remaining themselves, have both p_1 and p_2 . This certainly seems to make the distinctness from one another, of distinct conversational contexts, logically posterior to the difference, in the case of these property bearers, between their accidental and their essential properties. But if so, essential status cannot be *merely* the projection of what it is correct to say in the various conversational contexts.

1.5 Escape from Paradox

Conventionalism, I contend, ultimately founders on its refusal to allow that any objects in the world possess mind-independent existences. On pain of paradox we must allow that at least human minds themselves have mind-independent existences. Almost certainly we must also allow that human brains and bodies have mind-independent existences, and that the various material objects with which we interact have such existences as well.

But to make out these claims we must hold that the *essentialness* of the properties essential to nature’s kinds is independent of us—not a status for which we are responsible. And this returns us to the epistemological question: how do we manage to detect the essentialness of nature’s essential properties?