5. Essence and Natural Kinds: When Science Meets Preschooler Intuition[1]

Sarah-Jane Leslie

INTRODUCTION

It is common practice in philosophy to "rely on intuitions" in the course of an argument, or sometimes simply to establish a conclusion. One question that is therefore important to settle is: what is the source of these intuitions? Correspondingly: what is their epistemological status? Philosophical discussion often proceeds as though these intuitions stem from insight into the nature of things—as though they are born of rational reflection and judicious discernment. If these intuitions do not have some such status, then their role in philosophical theorizing rapidly becomes suspect. We would not, for example, wish to place philosophical weight on intuitions that are in effect the unreflective articulation of inchoate cognitive biases.

Developmental psychology has discovered a range of belief sets that emerge in the first few years of life, and which plausibly go beyond the evidence to which the child has had access in that time period. In such cases, it is reasonable to suppose that the belief sets do not derive solely from the child's rational reflection on her evidence, but rather show something about the way human beings are fundamentally disposed to see the world. (In some cases, the deep-seated dispositions are also shared with non-human animals.) There are many explanations of why we may be fundamentally disposed to see the world in a particular way, only one of which is that *metaphysically or scientifically speaking, the world actually is that way*. Another explanation may be that it is simply useful and practical to see the world that way—and this may be so even if it misleads us with respect to the metaphysical and scientific structure of reality. One particular way of carving up the world may be, say, efficient from the information processing point of view, without reflecting much about the fundamental nature of reality.

Suppose, then, we find that a particular set of philosophical intuitions closely resembles such an early-emerging implicit belief set. This certainly does not establish the falsity of the intuitions, but it should give us reason to subject them to further scrutiny. This is because such a set of intuitions

[1] I am indebted to Andrei Cimpian, Michael Devitt, Susan Gelman, Sally Haslanger, Elizabeth Harman, Mark Johnston, Paul Needham, Michael Weisberg, and an anonymous reviewer for helpful discussion, guidance, and invaluable feedback on earlier drafts. All the remaining mistakes are my own.

may strike us as completely compelling, and yet in fact reflect no more than a useful but ultimately misleading cognitive bias. Of course it may turn out that this is not the case; there is no argument from *early developing* to *false*. The point is purely epistemological: such a discovery should lead us to scrutinize the intuitions, and look for independent and converging evidence for the conclusions they urge.[2]

The present paper focuses on essentialism about natural kinds as a case study in order to illustrate this more general point. Saul Kripke and Hilary Putnam famously argued that natural kinds have essences, which are discovered by science, and which determine the extensions of our natural kind terms and concepts. This line of thought has been enormously influential in philosophy, and is often taken to have been established beyond doubt. The argument for the conclusion, however, makes critical use of intuitions, and I note that the intuitions are of the sort had by preschool children, and that they are traceable to a deep-seated cognitive outlook, which is often called "psychological essentialism." Further, if we did not have such a cognitive outlook or implicit belief set—a belief set which is in fact in place by at least age 4—then we would not have the relevant philosophical intuitions. In light of this, I consider the question of whether natural kinds actually have scientifically discovered or discoverable essences, and whether these putative essences could determine the extensions of our terms and concepts as Kripke, Putnam, and many others have supposed. In fact, a number of philosophers of biology and chemistry have argued that biological and chemical kinds do not have such essences, yet these arguments—particularly in the case of chemistry—have not been assimilated by philosophers more generally. The reason for this poor assimilation is, I suggest, that the Kripke/Putnam view *is just so intuitive*. But this fact, I argue, is due to a deep-seated cognitive bias, rather than to any special insight into the nature of reality.

The first half of the paper lays out the theory of psychological essentialism (which I rename *quintessentialism* for ease of exposition) in comprehensive detail, and summarizes some of the major experimental results that support it. I then consider further experimental evidence that suggests that lay people frequently misunderstand or misconstrue scientific findings as confirming their (quint)essentialist beliefs, when in fact the science does the opposite. Some of the more vivid and readily accessible cases involve quintessentialist beliefs concerning social kinds such as race and gender, and so such examples are also considered throughout the paper.

The second half of the paper takes up philosophical essentialism about natural kinds, and argues that quintessentialist beliefs are required for the crucial Kripke/Putnam intuitions. I then review extensive findings in biology

[2] The most general question here is the epistemological status of aspects of philosophical methodology, in light of recent psychological findings. Several theorists have recently begun to consider this question; for intriguing further discussion see, e.g. Gendler (2007), Stich (2009), and Weinberg, Nichols, and Stich (2001).

and chemistry and argue—along with a number of philosophers of biology and chemistry—that the findings conflict with Kripke/Putnam-style essentialism. To suppose otherwise is to misconstrue science through the lens of quintessentialism. The intuition that such essences *must* exist is no more than an expression of a belief set that is firmly in place by the middle of the preschool years, but this belief set does not comport with the complexities of actual science.

To drive this last point home, we must first understand the hypothesis of psychological essentialism and its detailed predictions, then recall the Kripke/Putnam view, and then investigate the actual facts as they stand in biology and chemistry. The psychology explains where our kind-essentialist intuitions come from, and the biology and chemistry explains why we should suspect and/or suspend those intuitions.

1. PSYCHOLOGICAL (QUINT)ESSENTIALISM

1.1. The Quintessentialists

Consider an intelligent species—let us dub them the Quintessentialists—whose members implicitly believe that a wide range of entities have within them a substance-like *quintessence*, which causally grounds their most important, stable, and enduring properties. Not all entities are believed to have quintessences: for example, artificial and manufactured items are rarely believed to have quintessences. However, biological beings, along with certain non-biological substances, are believed to be bearers of quintessence. It must be emphasized that the Quintessentialists rarely, if ever, *explicitly* entertain thoughts about quintessence as such; rather, their quintessentialist beliefs are tacit or implicit, though they are frequently manifested in a number of explicit ways.

Though they are not the only bearers of quintessence, biological individuals are paradigmatic bearers of quintessence. According to the quintessentialist belief set, each such individual has its own particular quintessence, yet there can be considerable similarities (and in the extreme case, potentially qualitative identity) between the quintessences of distinct individuals.[3] For example,

[3] Sometimes quintessentialism (= psychological essentialism) is characterized as not allowing for individual variation in quintessence—that is, sometimes it is assumed that members of the same natural kind will have the *same* quintessence rather than highly similar quintessence. I think in the general case, quintessentialism is best articulated as allowing at least for the possibility of individual variation. This is most vivid in the case of individual people; since a wide range of social kinds are quintessentialized, there must be as much variation between people as there are different quintessentialized social kinds for them to belong to. Further, the literature on organ transplantation (discussed in more detail in the section on the transmissibility of quintessence) reflects the belief that a recipient can take on individual characteristics of the donor—for example, heart recipient Claire Sylvia (1997) believed that she came to enjoy fried chicken and beer as the result of a transplant from a young man named Tim (see also Inspector, Kutz, and David, 2004). In order to accommodate these sorts of considerations, the possibility

offspring have quintessences that are highly similar to those of their parents. Further, two individuals are generally only able (or disposed) to produce offspring if their quintessences are sufficiently alike. Like begets like, as far as quintessence is concerned, and Quintessentialists are convinced that one's parents' quintessence is the main determiner of one's own quintessence. These sorts of similarities in quintessence are taken to form the basis of real or "natural" kinds, as opposed to merely conventional groupings. (In the case of non-biological individuals, such as certain substances, origins may also be considered important—for example, if two substances are formed in very different ways, then it is reasonable to suppose they have distinct quintessences.)

Thus the Quintessentialists believe that there are ways of dividing individuals up into kinds that carve nature at its joints—namely the ones that group according to objective likeness in quintessence. Conversely, they also acknowledges that some groupings are merely nominal, and do not reflect genuine similarity classes of quintessences. However, Quintessentialists firmly believe that many of their own words and concepts map directly on to the real, objective kinds.

The Quintessentialists also believe that there are a number of levels or degrees of similarity in quintessence, all of which are real and objective, in effect constituting a taxonomic hierarchy of kinds. At the lowest levels of this taxonomy, there are considerable similarities between the quintessences of members of some distinct kinds, while at the higher levels, there is considerable variation between the quintessences of members of the same kind. Importantly, there is a *privileged* level of this subjective taxonomy that occupies a "sweet spot" in this trade-off between within-kind variation in quintessence, and cross-kind quintessential distinctness. At this level, individual members of the same kind have only minimal differences in their quintessences, and these quintessences are quite dramatically different from the quintessences had by members of other kinds. (The Quintessentialists' cognitive psychologists call this taxonomic level "the basic-level.") The Quintessentialists believe that this privileged taxonomic level is objectively determined, and so there is a privileged way of answering the question of whether a given individual is the *same kind of animal* (or *same kind of plant*) as another: namely whether they belong to the same basic-level kind. The notion of a basic-level kind extends beyond the biological realm and into the chemical realm too; samples of substances can also share quintessences, and there is again a privileged level of the taxonomy of substances where the kinds in play maximize intra-kind similarity (in the limit case, intra-kind qualitative identity) and inter-kind

should exist for individual variation in quintessence. It should further be emphasized that the model I am introducing here would allow also for the possibility that, in some cases, the nature of the individual's quintessence would be exhausted by the nature of the quintessence of the (most specific) natural kind to which they belong. (That is, my model is compatible with the limit case where similarity becomes identity.) For example, one might believe that, e.g. mosquitoes do not display variation in their individual quintessences.

dissimilarity in quintessence. Thus the Quintessentialists are strongly inclined to believe there is a privileged way of answering the question of whether one thing is the *same substance* as another.

It must be emphasized that the Quintessentialists believe that these taxonomic categories are determined by objective facts about quintessence. This has several consequences; firstly that the Quintessentialists strongly believe that membership in these kinds is a purely intrinsic matter—a given individual's quintessence is the sole determiner of its membership in a real kind. Further, they believe that quintessence lends itself to being "carved at its joints"—that is, quintessence does not vary continuously between individuals of different kinds, but rather is objectively distributed in such a way that, especially at the basic-level, members of the same kind have considerable sameness of quintessence, while non-members have distinctly different quintessences. Thus, membership in these kinds ought to be close to an all-or-nothing matter; that is, Quintessentialists believe that real kinds should have sharp boundaries. Sometimes they themselves are not able to tell for certain the kind to which an individual belongs, but they believe there should be a correct answer to the question—an answer that is determined by the individual's quintessence. (However, other beliefs about quintessence—namely that it is transmittable and mixable, as discussed below—compete with this one, creating an inevitable tension.)

The single most important feature of quintessences is that they have causal powers. An individual's quintessence is understood to be the causal root of many/most of that individual's stable and enduring intrinsic properties. The respects in which members of a real kind are outwardly similar to each other is understood by the Quintessentialists to be the direct upshot of the similarities between their quintessences. Thus, the Quintessentialists consider outward signs of kind-membership to be important, but only in as much as they are indicators of underlying similarities in quintessence. It would be a mistake, though, to think that quintessences *only* have kind-related causal powers. For example, the Quintessentialists believe that they themselves each have a distinct quintessence, which grounds many of their individually distinguishing properties (in addition, that is, to grounding kind-general properties). For example, individual personality traits are considered to be the upshot of the individual's quintessence.[4]

The Quintessentialists further believe that an item's quintessence causes its properties in a defeasible way—that is, adventitious factors can prevent an individual from manifesting all the properties that the quintessence would otherwise cause it to manifest. This is particularly clear to them when they consider kind-wide quintessential properties that have been altered by clearly external circumstances—for example, some Quintessentialists have

[4] Again, not all articulations of quintessentialism (= psychological essentialism) take this line, however it seems to me the best way to account for a number of findings; see footnote 2, and the section below on the transmissibility of quintessence.

lost limbs to amputation or accident. It can also happen that an individual's quintessence is somehow thwarted from ever expressing itself in a particular feature—for example, some Quintessentialists are simply born without limbs.

Since quintessence plays such a central role in beliefs about kind-membership, and since quintessence causally grounds a range of properties, Quintessentialists are ready and able to exploit kind-membership in their reasoning practices. Since natural kinds—especially "basic-level" kinds—are supposed to be groups with highly similar quintessences, Quintessentialists treat membership in a natural kind as a cornerstone of their inductive practices. If one member of a basic-level kind has a plausibly quintessential property, then other members of that kind may well have the property too. (Of course, these inferences are defeasible in a number of ways, given the nature of quintessential properties. They are, nonetheless, a good default starting point.) In this way, knowledge of kind-membership is more important to the Quintessentialists' inductive inferences than anything else: more important than perceptual similarity between individuals, and more important than their sharing any other properties.

When it comes to beliefs about their fellows, the same framework is employed. Quintessentialists manifest a tendency to believe that certain social categories—most prominently race and gender categories—group according to real similarities and distinctions in quintessence, and so reflect objective, non-constructed differences in nature. Just as in the non-social case, quintessentialist beliefs about these categories involve thinking that membership in them is determined by natural, inborn, intrinsic properties of the individual; that members of these groups are highly similar to one another and distinctly different from non-members; and that membership in these groups is more important and informative about the person than his or her appearance, demeanor, and other such readily observable properties. This quintessentialist belief set about the social domain has a number of unfortunate consequences; in particular, social groups whose members are seen to share very similar quintessences often experience significant degrees of social prejudice.

When it comes to the nature of quintessence itself, the Quintessentialists do not have very specific beliefs at all. Instead, they have what might be termed a 'placeholder' notion of quintessence—they are sure that *something* fills this role, but they know not what. They are confident that it has a nature and independent existence, however, and they see no reason why they shouldn't discover more about its nature, especially as their scientific practices mature. However, if one were to push for an elaboration of the metaphysics of quintessence, one might arrive at something like the following: quintessences are substance-like entities in that they occupy—or better, pervade—space-time regions. They can also mix with each other, albeit only under unusual circumstances. Quintessences can permeate objects, including inanimate objects, and thus sometimes be transmitted. Notably, such transmissions almost always require a physical bearer.

In the normal case, an individual's quintessence pervades its insides; internal parts are regarded as more important for quintessence than external surface parts. In this way, Quintessentialists believe that altering an individual's insides is more likely to produce a change in the individual's quintessence than altering an individual's surface properties. In keeping with their beliefs that real, basic-level kinds reflect highly similar quintessences, they also expect that members of the same basic-level kind will have highly similar insides—even if their external appearances are quite different.

Since one's insides are suffused with one's quintessence, the possibility of physically transplanting a part of an individual's insides into another raises the possibility of transferring some of an individual's quintessence into another. In recent years, the Quintessentialists' medical technology has evolved to a point where this is more than just a possibility: internal organs can now be transplanted from one Quintessentialist to another. This raises particular anxieties for the Quintessentialists, since they believe that this practice transfers some of the donor's quintessence to the recipient. The donor's quintessence retains its causal powers, and as a result, the Quintessentialists believe that the recipient may be irreversibly changed. (The Quintessentialists are often reminded by their medical professionals that there is no scientific evidence to support this claim, but the lay belief persists nonetheless.)

Recently, some Quintessentialist scientists have made considerable technological progress on the possibility of transplanting organs from other species. This is most disquieting to the Quintessentialist public. They fear that receiving transplants from other species may alter them, making them more animal-like. They feel that mixing the quintessences of different basic-level kinds is profoundly unnatural.

While internal parts are the main bearers of quintessence, it strikes the Quintessentialists that a small amount of an individual's quintessence 'rubs off' on things that he or she handles and uses. This belief manifests itself in several ways. For example, Quintessentialists are willing to pay large sums of money for otherwise unremarkable objects, just because they were once owned and used by a celebrated person. They also shrink from items that were handled or used by an infamous and hated person. (This disposition increases as the contact becomes more intimate—e.g. a worn item of clothing is more potent in this respect than a touched pen.)

At least to some extent, the Quintessentialists believe that they can be causally affected by quintessence transmitted through an object in this way. To the extent that they hold this belief, it may be weaker than their belief in the potency of organ transplantation—which is, of course, reasonable given their beliefs about the relationship between internal parts and quintessence. There is, nonetheless, some degree of belief in the causal potency of quintessence transmitted even in such an indirect way.

It must be emphasized again that these beliefs in quintessence are rarely *explicitly* formulated, articulated, or entertained by the Quintessentialists. Rather, the belief structure sketched out here is a tacit framework that guides their thinking. Thus, to determine whether an individual is a Quintessentialist, it does not do to ask them directly—rather, one must use a number of indirect measures. Further, the Quintessentialists are not *taught* to have the beliefs they do. These beliefs are not a mere cultural artifact: Quintessentialists from a wide variety of culturally independent groups exhibit the belief structure sketched above. And tellingly, Quintessentialist children manifest this way of thinking from a very young age. The role of culture is often to fill out more specific forms of the implicit beliefs that are characteristic of the early-developing quintessentialist syndrome.

1.2. Psychological (Quint)Essentialism

Psychological essentialism (henceforth *psychological quintessentialism*) holds that we are the Quintessentialists described above. This well-confirmed psychological hypothesis attributes to us an implicit and early-developing belief in quintessence—in a substance-like entity possessed by some individuals and some forms of stuff, a substance which pervades their insides and causally grounds many of their stable and persisting properties. Membership in a natural kind is taken to be determined by objective features of one's quintessence—namely the features that are shared by other members of the kind.

It is crucial to understand the hypothesis of psychological quintessentialism correctly. It does not involve the claim that to be a Quintessentialist is to subscribe to *all* the beliefs listed above; rather quintessentialism is better thought of as a syndrome, which can be manifested in a variety of default implicit beliefs and ways of interpreting one's world. Thus, one might be a Quintessentialist without holding each belief mentioned above. Compare, for example, the symptoms of the syndrome that is depression. Such symptoms include *having suicidal thoughts*; however it would be a mistake to suppose that one simply could not be depressed unless one has suicidal thoughts.

Moreover, default implicit beliefs *can be* suppressed by explicit learning of contrary facts. For example, a well-informed adult may be less likely to explicitly hold the relevant beliefs about, say, organ transplantation. But this is fully compatible with being a Quintessentialist; that is, exhibiting the psychological syndrome that is manifested in the default beliefs mentioned above.

It would also be a mistake to suppose that introspection alone could deliver the result that one is not a Quintessentialist. Introspection at best provides access to *explicitly* held beliefs; however, it is unhelpful in assessing beliefs that are largely implicit. Sometimes, one belief may be held implicitly, while a contradictory belief may be held explicitly. A dramatic illustration of this phenomenon concerns prejudiced attitudes as revealed by tests such as the Implicit Association Test (IAT) (e.g., Greenwald and Banaji, 1995). Many individuals who honestly explicitly disavow sexist and racist (and other prejudiced) beliefs nonetheless may implicitly hold such attitudes—even dedicated feminists may implicitly associate, for example women with nurturance but men with intelligence. Further, implicit attitudes have been found to better predict a range of behaviors as compared to explicit/introspection-based measures of attitudes (Greenwald, Poehlman, Uhlmann, and Banaji, 2009; for excellent further philosophical discussion of these points and related ones, see Gendler, 2008a, 2008b, in press). Implicit attitudes and beliefs are not somehow of lesser significance than explicit ones—only a limited range of the beliefs and attitudes that shape our behavior are available to introspection. And again, the explicit rejection of a given belief does not entail that that belief is not held implicitly.

So quintessentialism is a syndrome, which encourages *one or another* cluster or range of default implicit beliefs drawn from the inventory provided in the previous section. These beliefs are not readily available to introspection, and they can persist in the face of known evidence to the contrary. The combination of these three observations means that a certain tempting response to the arguments given in this paper would be mistaken. I have in mind the following protestation, crudely characterized: "I know that it is scientifically impossible for a heart transplant to alter one's personality, therefore I am not a Quintessentialist; however I still have the Kripke/Putnam intuitions, therefore it cannot be the case that Kripke/Putnam intuitions are due to quintessentialism!" This line of objection, along with its more subtle variants, turns on a failure to understand the hypothesis of psychological quintessentialism, and the significance of implicit beliefs more generally.

1.2.1. Some Crucial Findings

Despite the fact that the beliefs in question are largely implicit, and so are somewhat more difficult to discern than our explicit beliefs, the last two decades or so have produced a wealth of empirical evidence suggesting we are, indeed, Quintessentialists. In the interests of space, I will not attempt a complete review here, but will rather highlight some of the main findings. (For an extremely thorough review of the data available through 2003, see Gelman, 2003.)

There is a considerable amount of evidence that supports the claim that there is a privileged level in our psychological taxonomies—namely the so-called "basic-level" (Rosch, 1978; see also Coley, Medin, and Atran, 1997; Coley, Medin, Proffitt, Lynch, and Atran, 1999; Gelman, 2003; and many others). Members of the same basic-level kind are perceived to have quintessences that are highly similar, and also highly distinctive, in that members of other basic-level kinds have notably different quintessences (Gelman, 2003). Accordingly, in inductive reasoning tasks, both children and adults very often generalize properties to other members of the same basic-level kind (e.g. tigers), but not to more inclusive kinds (e.g. mammals); nor do they limit their generalizations to less inclusive kinds (e.g. Bengal tigers) (Coley et al., 1997; Gelman and O'Reilly, 1988; Waxman, Lynch, Casey, and Baer, 1997).

Information about membership in a basic-level kind has a powerful impact on people's inductive inferences from a very early age. For example, even two-year-olds prefer to base their inductive inferences on information about kind membership (as communicated via language) than on perceptual similarity (Gelman and Coley, 1990)—despite the general importance of perceptual similarity (especially shape) in guiding young children's inferences. Further, preschool children distinguish between the sorts of properties that can be generalized along kind boundaries from those that cannot: children do not generalize transient or accidental properties such *having fallen on the floor that morning* (Gelman and Coley, 1990; Gelman, Collman, and Maccoby, 1986; Gelman and Markman, 1986, 1987). The latter are paradigmatically the sorts of properties that are independent of one's quintessence.

It is also worth noting that the results of many of the studies cited above depend on even very young children being willing to accept that an individual can appear for all the world like a member of one kind, yet *really* belong to another kind. Further, it is kind membership that is understood to be most important when making inferences about deep, shared properties. Notably, these findings extend to substance kinds, with preschoolers understanding that something can look more like coal than gold yet still be gold, and also basing their inferences about projectible, non-obvious properties on this shared kind membership (Gelman and Markman, 1986).

Young children also have definite views on the causal roles of nature vs nurture. For example, preschoolers believe that an infant creature with kangaroo parents will grow up have a pouch and hop, even if it is raised exclusively with goats (Gelman and Wellman, 1991; see also Heyman and Gelman, 2000; Solomon and Johnson, 2000; Springer and Keil, 1989). Young children evidence the same set of beliefs when it comes to gender: for example, that a male baby raised on an island populated only by females will grow up to display boy-typical behavior and preferences (e.g. preferring to play with trucks rather than dolls; Taylor, 1996; Taylor, Rhodes and Gelman, 2009). Preschoolers also display at least some nature-over-nurture beliefs in the case of race, expecting that a child born to black parents will grow up to be black (Hirschfeld, 1996), though in general quintessentialist beliefs about race develop later than quintessentialist beliefs about species and gender (e.g. Rhodes and Gelman, 2009a).

More generally, a number of studies have documented beliefs in children from very different cultural backgrounds about the relative power of nature vs nurture, for example Menominee Indian children, Yukatek-Mayan children, and urban Brazilian children (Atran, Medin, Lynch, Vapnarsky, Ek' and Sousa, 2001; Sousa, Atran, and Medin, 2002; Waxman, Medin and Ross, 2007). These findings are exactly what one would expect if one's subjects were Quintessentialists.

Children and adults also tend to 'intensify the boundaries' of natural kinds—that is, they judge that there should be sharp, exclusive boundaries between natural kinds (at the same taxonomic level). For example, Diesendruck and Gelman (1999) found that adults almost always judged that a given animal was definitely a member of a kind or definitely not a member of a kind; their judgments of how typical an exemplar of the kind the individual was, however, was more graded. The same pattern did not hold for artifact kinds, however, reflecting that this pattern is specific to natural kinds. Rhodes and Gelman (2009b) documented similar beliefs in young children. Further, even if adults are themselves uncertain as to how to classify an animal, they judge that there is a single correct answer which an expert could determine (Luhman and Coley, 2000). Young children also judge that membership in both animal and gender kinds reflect correct, objective, non-conventional facts (Kalish, 1998; Rhodes and Gelman, 2009a).

Children at least as young as four understand that animate individuals have very different insides from inanimate individuals (e.g. Gelman, 1987; Gelman and O'Reilly, 1988; Gelman and Wellman, 1991; Gottfried and Gelman, 2005; Simons and Keil, 1995). They also understand that the internal parts of animate creatures have special importance when it comes to determining both an individual's behavior and its kind-membership (Gelman and Wellman, 1991). Some recent work on infants suggests that the importance of insides is recognized from a very young age. Newman, Herrmann, Wynn, and Keil (2008) found that fourteen-month-olds expected that animals with the same internal features (e.g. visible red stomachs), rather than external features (e.g. blue hats), would display the same sorts of self-generated motion. However, when the motion of the target did not appear to be self-generated, these expectations did not arise. Thus it appears that, when it comes to apparently animate, self-generated patterns of motion, even infants expect that internal parts will be better predictors than external parts.

Additional studies have found that both children and adults understand that insides are of considerable importance in determining membership in both biological and chemical kinds (e.g. Keil, 1989; Lizotte and Gelman, 1999; Newman and Keil, 2008). For example, Keil (1989) reports a classic series of experiments that show children's increasingly sophisticated understanding of the importance of insides in determining kind-membership, with even very young children judging that, for example, a raccoon that has had its fur shaved and dyed so as to make it look just like a skunk is nonetheless still a raccoon. Interestingly, recent evidence suggests that rhesus monkeys also privilege insides over outward appearance in determining kind-membership (Phillips, Shankar, and Santos, 2010; see also Phillips and Santos, 2007).

The majority of studies mentioned so far were conducted with participants primarily from North American communities. However, recent work has discovered similar dispositions among other cultural groups. As of now, essentialist beliefs have been documented in at least the following communities: Indians (Mahalingam, 2003; Meyer, Leslie, Gelman, and Stilwell, in press), Brazilians (Diesendruck, 2001; Sousa et al.,2002), the Menominee (Waxman et al., 2007), the Vezo in Madagascar (Astuti, Solomon, and Carey, 2004), Yucatec Mayans (Atran et al., 2001), the Yoruba in Nigeria (Walker, 1999), and the Torguud of Mongolia (Gil-White, 2001). Thus, while of course more work is needed, the available evidence suggests that quintessentialist thinking is not a local phenomenon, but rather is a pervasive aspect of human psychology.

1.2.2. The Transmissibility of Quintessence

One question that arises at this point is why quintessence should be modeled as something substance-like, transmissible, and mixable; the above data do not obviously require this to be so. Consider, however, Sir Arthur Conan Doyle's *Adventure of the Creeping Man*. In this story, Sherlock Holmes is called to investigate some alarming and mysterious goings-on at the house of an ageing professor. Through a characteristically brilliant string of 'deductions', Holmes figures out what is happening: in an effort to recapture his youthful vitality, the professor has been injecting himself with a serum extracted from the glands of monkeys. This has caused the professor to take on several monkey-like characteristics: he has become increasingly aggressive and his knuckles have become thickened and hairy; he has developed superhuman climbing abilities; and at times he adopts the gait of a monkey.

This story—however much at odds with modern scientific understanding—reflects a basic quintessentialist way of thinking. A serum extracted from the glands of a monkey confers a portion of the monkey's quintessence, complete with its causal powers, thus making its recipient take on the characteristics of the monkey. The tale reflects some of the most specific characteristics of quintessentialist thinking: the notion that the quintessence is causally potent, lies in the internal parts of the individual, and is substance-like in that it can be physically removed, relocated, and mixed with the recipient's own quintessence to produce a hybrid mix of outward properties. (Does this explain the demand for bulls' testicles among those who desire a certain 'outward' property?)

Can such beliefs be systematically studied and documented? In an early study, Johnson (1990) found evidence that young children do indeed believe that such hybrid mixtures of quintessence are possible; for example, they believe that a heart transplant from a mean person would make the recipient become meaner. Meredith Meyer, Susan Gelman, and I replicated and extended these findings, and found they held up even when children were asked to consider characteristics that are not culturally associated with hearts, for example *being smart* (Meyer, Gelman, and Leslie, submitted). In collaboration with Sarah Stilwell, we also recently conducted a series of experiments to test for such beliefs in adults. In particular, we wished to assess whether people would endorse the possibility of a transplant or transfusion recipient's personality changing to become more like the donor's. We asked participants to consider donors from a range of social categories and two animal categories. We studied adult participants in both the United States and India, and have found that a sizeable portion of participants from both communities judge that the recipient's personality may change to become more like the donor's as a result of receiving a transplant or transfusion (Meyer, Leslie, Gelman, and Stilwell, in press).[5]

Are transplants and transfusions the only means by which quintessence may be transmitted? A quick perusal of auction catalogs will show that people are willing to pay large sums of money for rather ordinary objects, if those objects once had contact with a famous or admired person. Hood and Bloom (2008) and Frazier and Gelman (2009) found that even young children place higher value on such objects, as compared to qualitatively identical duplicate objects. The converse of valuing objects that have been in contact with admired persons is shrinking from items that have been in contact with reviled individuals; this phenomenon has been studied and documented at length by Carol Nemeroff, Paul Rozin, and their colleagues (see Rozin and Nemeroff (2002) for a review). For example, in one famous study, Nemeroff and Rozin (1994) asked participants to rate how they would feel about wearing a sweater that had come into contact with various people, desirable and undesirable, including an enemy of theirs, and a person that the participants judged to be evil. (Unsurprisingly, Adolf Hitler was the favored choice for the latter category.) Participants were strongly opposed to wearing a sweater worn by either a personal enemy or by Hitler; in fact they rated a sweater that had fallen in dog feces as more desirable to wear than Hitler's sweater.

Notice that such results are not accounted for by the simple observation that we happen to value or care about historical properties. There are many historical properties of paintings and sweaters and keepsakes and the like that we do not care about, such as having once been located in Europe or twice having been scanned by an x-ray machine. The crucial thing to be explained is the quasi-infection-based character of the "authenticity and contamination" findings, and a natural hypothesis is that it is an implicit belief in transmissible quintessence which accounts for this.

As with transplants and transfusions, the clearest way to determine whether the "authenticity and contamination" findings are evidence of a belief in a transmissible quintessence would be to see whether people believe that wear-

[5] Such beliefs are also attested to in a number of (non-experimental) surveys of attitudes towards transplants (Basch, 1973; Belk, 1990;Hayward and Madill, 2003; Inspector et al., 2004; Sanner, 2001a, 2001b). For example, Inspector et al (2004) found that a third of heart-transplant recipients believed they had in fact taken on characteristics of the donor. Sanner (2001a, b) interviewed people about their attitudes towards various organ transplants, and she reports participants' explicit formulations of 'Creeping Man' concerns:
'I would perhaps look more piggish with a pig's kidney.'.
'Would I become half a pig, if I got an organ from a pig?'
'What if I would start grunting?' (2001a, p. 22).
A number of participants also expressed more general discomfort at the idea of 'mixing' species in this way:
'I feel instinctively that it's wrong to mix different species, it would go wrong.' 'My body would let me know that an animal organ didn't fit. It's contrary to nature.'
'It's unnatural to move body parts between species' (2001b, p. 1495).
And quite decisively:
' The whole pig nature just feels like a big no' (2001b, p.1496).

ing an item of clothing might cause a person to become more like the original wearer, even if the person in question does not know about the item's history. For example, suppose one is about to stand trial, and one learns that the judge is unwittingly wearing an item of Stalin's clothing, worn by him on and off over several years. Might one become anxious that the judge will be less just as a result? This phenomenon has not been extensively studied; however, one experiment suggests that such beliefs can be found. In particular, Johnson and Jacobs (2001) found that elementary school children and adults alike believed that wearing Mr. Rogers' sweater would make a person friendlier, even if the wearer did not know about the prior owner. This finding suggests that, just as with transplants and transfusions, the notion of transmissible quintessence is at play here. The history of close-range adoration of relics purportedly from sainted persons may also be worthy of examination.

1.3. Quintessentialism and Science

Some quintessentialist beliefs—such as the beliefs that receiving a blood transfusion or that wearing Mr. Roger's sweater can alter one's personality—are quite clearly at odds with the deliverances of science. Many other quintessentialist beliefs are, I would argue, similarly at odds with science, although on the surface they do not seem so. In particular, scientific findings, particularly in biology, are often misunderstood by the general population in a specifically quintessentialist manner. That is, people inappropriately map scientific concepts onto their pre-existing quintessentialist beliefs, and then consider those beliefs to be scientifically underwritten. (This may be part of a more general phenomenon, whereby people adopt culturally available ways of elaborating and articulating the same core, quintessentialist beliefs (e.g. Waxman et al., 2007).)

While most educated adults have some familiarity with the concepts of genes and of DNA, misunderstandings abound, many of which at least indirectly suggest that people may simply have mapped the concepts onto their implicit notion of quintessence. One illustrative pair of anecdotal examples can be drawn from the geopolitical sphere: on 13 March 2000, Vladimir Putin reportedly asserted that "central power is in Russia's genes."[6] Apparently President G. W. Bush concurred; at a White House news conference held on 17 October 2007, he wondered "whether or not it's possible to reprogram the kind of basic Russian DNA, which is a centralized authority."[7] There is, of course, no such thing as specifically Russian DNA; however a quintessentialist view of Russians as a social kind entails belief in a particular Russian quintessence. There is also, of course, no scientific evidence to suggest that genetics per se would specially predispose a person or a group of people to

[6] Oleg Shchedrov, "Central Power in *Russia's Genes*, *Putin* Says," Reuters (Moscow), 13 March 2000, from Johnson's Russia List, 13 March 2000.

[7] "Russia's DNA," *Washington Post* editorial, 19 October 2007.

accept centralized authority, yet quintessentialist thinking is highly compatible with such a belief.[8]

As the above example illustrates, "DNA" and "genes" are often invoked in people's conceptions of quintessentialized social groups. A number of social groups, such as groups demarcated by race and gender, are often highly quintessentialized; that is, people believe that the members of those social groups have highly similar quintessences to one another, and further that non-members have quite distinct quintessences—in effect, these groups are treated as social analogs of basic-level kinds. Lay beliefs about race and genetics often show precisely this pattern. For example, the majority of adults in the United States agree with the following statement: "Two people from the same race will always be more genetically similar to each other than two people from different races" (Jayaratne, 2001). Such a belief is wholly at odds with contemporary scientific thinking about genetic variability. Instead, the genetic variability *within* a racial group is just as high as the degree of variability *across* racial groups (e.g. Graves 2001; Lewontin, Rose, and Kamin, 1984; Templeton, 1999). In general, contemporary genetics is far more concerned with understanding how genetic differences among *individuals* can explain the phenotypic differences among those individuals, rather than attempting to find genetic explanations for putative group differences. This point has been poorly assimilated, however. As an illustration, when asked to consider possession of the trait *being nurturing*, Cole, Jayaratne, Cecchi, Feldbaum, and Petty (2007) found that people thought that genetics could *better* explain (perceived) gender differences than individual differences within genders. That is, contrary to the actual deliverances of genetics, they thought that group-level genetic explanations were more applicable than individual-level ones.[9]

Social groups aside, there are numerous other examples of quintessentialist misinterpretations of genetics. For example, Lanie, Jayaratne, Sheldon, Kardia, and Petty (2004) note that people with a family history of a heritable disease sometimes think that they can't get the disease if they don't look like the family members who have it. This is a natural belief to have against the backdrop of quintessentialism: one is susceptible to a heritable disease only in as much as one shares quintessence with the relevant relative, and similarity of appearance is an excellent guide to the extent to which quintessence is shared. People also confidently attribute traits, abilities, and dispositions to their genes in the absence of any scientific evidence supporting the idea that those characteristics are genetically based. Lanie et al. report that, even though few respondents could give a remotely adequate characterization of what genes even are, "it is interesting that elsewhere in the interview about three quarters of respondents had no trouble giving an example of at least one nonphysical,

[8] I am extremely grateful to Peter Godrey-Smith for drawing my attention to these examples.

[9] Interestingly, Keller (2005) found that belief in "genetic determinism"—that one's genes are a powerful determiner of one's traits and character—correlated with social prejudice, and that priming people to think about genetics further increased their level of prejudice.

nonmedical 'genetic' characteristic that ran in their families, even if there was no scientific research in support of their beliefs" (2004, p. 311). Such examples included characteristics like *being good at home repairs* (and perhaps also could have included *embracing centralized authority*).

As another illustration, consider that on the quintessentialist outlook, a quintessence has the power to cause its bearer to have certain characteristics, more or less independently of the bearer's environment. And so relatedly, the outlook encourages the belief that if a certain characteristic does not completely spring from the bearer's quintessence, then it must be determined by the environment. Thus, quintessence and environment are seen as largely mutually exclusive influences when it comes to determining the causal source of a trait.[10] In contrast, modern genetics holds that phenotypic traits arise from complex *interactions* between an individual's genes and its environment. Genetics and environment are not in any sense mutually exclusive as causes, rather they operate in tandem. Nevertheless, lay understanding of genetics often implements the quintessentialist model, in which genetic causes and environmental causes are seen as exclusive of one another. For example, in an extensive study, Jayaratne, Gelman, Feldbaum, Sheldon, Petty, and Kardia (2009) found numerous negative correlations, and no positive correlations, between participants' endorsements of genetic and environmental factors in explaining the source of a range of traits.

In recent years, a number of researchers have argued that quintessentialism is responsible for both resistance to and misunderstanding of evolutionary theory (e.g. Gelman and Rhodes, in press; Hull, 1965; Mayr, 1982, 1988, 1991; Samarapungavan and Wiers, 1997; Shtulman, 2006; Shtulman and Schulz, 2008). The main culprit is the quintessentialist belief that within-species variability is a limited and uninteresting phenomenon, in comparison to within-species similarity. That is, quintessentialism places considerable emphasis on the similarities between members of a given species (or more accurately, a given basic-level kind, since these are not always one and the same), and de-emphasizes the differences between them. (This is, of course, why quintessentialism encourages such rich inductive inferences over these basic-level kinds.) To correctly understand evolutionary theory, however, one must appreciate that there is a considerable degree of variability within a species, since this variability is *precisely what allows the evolutionary process to occur.*

Not only does quintessentialism make people resistant to accepting evolutionary theory—for example, Samarapungavan and Wiers (1997) found a sizeable portion of Dutch third- and fifth-graders believe the species to be 'eternal and unchanging'—but it further means that even educated adults often wholly misunderstand what the theory claims. Shtulman (2006) found

[10] Of course this is not intended as an absolute and universal quintessentialist principle, but rather a general guiding one—for example, there could still be cases in which the environment prevents a quintessential property from being had, an obvious example being a tiger that has lost its tail.

that the majority of the talented group of students enrolled in the Harvard Summer School held "transformational" rather than "variational" conceptions of evolutionary theory. The correct variational conception of evolution holds that random mutations occur all the time, and some of these confer an adaptive benefit on their possessors, and so render them more likely to successfully reproduce and pass along the mutation to the next generation. The transformational conception of evolution, in contrast, holds that, if a trait is beneficial to the species, then over generations the species as a whole is likely to acquire the trait. Thus, on the transformational view, the quintessentialist emphasis on within-species commonalities is preserved—what happens is that (somehow) over time the common quintessence of the species alters so as to produce more beneficial traits. Perhaps unsurprisingly, people who hold this transformational view of evolution are less likely to believe in evolution as scientific fact. After all, what they understand it to be saying is not in fact true.

As emerges at some length below, biology's understanding of species and other taxa is deeply at odds with the quintessentialist mindset. The relationship between one's genotype and one's species membership is complex, probabilistic, and highly dependent on extrinstic factors; there is no such thing as the 'species' genotype.

Just as the systematic misconstruals of evolutionary theory illustrate, educated adults may appear to have a working understanding of biology, while in fact their understanding is rife with quintessentialist confusion. For example, consider the following remark by the brilliant semanticists Hans Kamp and Barbara Partee: "the vast majority of natural kind terms are sharp in the strict sense of being determinately true or false of everything that is found in the real world. For instance, to belong to a particular biological species an individual must have the DNA of that species; and almost without exception this is a property which an individual organism either definitely has or else definitely lacks" (1995, p. 175).[11] How far does this kind of quintesssentialist thinking extend, and has it left its mark on philosophy?

2. PHILOSOPHICAL ESSENTIALISM AND PSYCHOLOGICAL QUINTESSENTIALISM

2.1. Natural Kinds and Philosophical Essentialism

Hilary Putnam and Saul Kripke famously argued for a version of philosophical essentialism as applied to natural kinds, according to which there are scientifically discoverable necessary and sufficient conditions for belonging to a natural kind. In particular, these necessary and sufficient conditions consist in a specification of the kind's hidden underlying structure. This underlying

[11] I am grateful to Susan Gelman for directing my attention to this passage. She and Marjorie Rhodes use this passage to illustrate a very similar point in Gelman and Rhodes (in press).

structure is discoverable by science, and plays a causal/explanatory role in the determination of the kind's perceptible, or manifest, properties. Both Kripke and Putnam further argue that the extensions of natural kind terms are determined by sameness of underlying essence—rather than by, for example, sameness of manifest properties.[12]

2.1.1. The Twin-Earth Thought Experiment

Perhaps the best-known illustration of the point is Putnam's Twin-Earth thought experiment. We are asked to imagine a planet that is qualitatively identical to Earth, except that where Earth contains H_2O, Twin-Earth contains a superficially indistinguishable substance whose complex chemical formula is abbreviated as "XYZ." Thus, on Twin-Earth, XYZ fills the oceans and lakes, and falls from the sky as rain; Twin-Earth people drink XYZ to quench their thirst, bathe in it, use to make soups, and so forth. Putnam then asks us to imagine that, in 1950, some inhabitants of Earth set out on a spaceship and reach Twin-Earth. When they arrive there, they are astonished at the similarities, including what they initially believe to be the abundance of water on the planet. Crucially, though, we are encouraged to have the intuition that when the Earth's astronauts say "there is water in the lakes on Twin Earth," what they say is false. If they go on to perform chemical tests on the liquid that fills the oceans and lakes on Twin Earth, they will realize they were just wrong to have called the substance "water." This is because the extension of *our* term "water" picks out all and only that which has the underlying chemical composition H_2O.

Alternatively, imagine an inhabitant of Earth, Oscar, and his Twin-Earthian doppelganger, Twin-Oscar. Putnam argues that the word "water" as used by each has different extensions—for Oscar "water" applies to H_2O, for Twin-Oscar "water" applies to XYZ. Of course, Oscar and Twin-Oscar may have different beliefs about water and twin-water respectively: namely, Oscar may believe that water is composed of H_2O, while Twin-Oscar may believe that twin-water is composed of XYZ. However, Putnam then asks us to 'roll the clock back' to 1750, before anything was known about the chemical compo-

[12] It must be acknowledged that Hilary Putnam came to revise his view on the matter, and allow that, e.g. our interests also figure in the determination of the extension of natural kind terms (e.g. Putnam, 1992). This revised view has not, however, been nearly so influential as his original view, and in fact comparatively few philosophers are even aware that Putnam changed his view (Hacking, 2007a). The discussion in this half of the paper is directed towards the view—articulated in "The Meaning of 'Meaning,'' at least as it is widely interpreted—that has been so influential, and which is so often characterized as Kripke/Putnam essentialism. For further discussion, see Hacking (2007a); nothing in this paper disagrees with Hacking on the historical observations, however I will continue to speak of the "Kripke/Putnam view," since, as Hacking himself observes, this is how the view is widely understood.

There is a crucial further point. It is one thing to qualify the view in the light of isolated counterexamples and quite another to hope that this will deal with the full systematic range of examples of natural kind terms drawn from biology and chemistry. What follows is intended to strongly suggest that we much more than the concession that our interests play a role.

sition of water. Putnam maintains that "water" still has a different extension in the mouths of Oscar and Twin-Oscar—even though we may stipulate now that neither Oscar nor Twin-Oscar (nor any local experts) have any beliefs about the nature of the chemical composition of these substances. If we agree with these intuitions, we must suppose that two people can be in duplicate inner psychological states, while their terms have different extensions, due to the nature of the local natural kinds. Putnam famously concludes that "meanings just ain't in the head!" (1975, p. 227).

In what follows below, I will argue that the intuitions associated with Putnam's highly influential thought experiment are quintessentialist intuitions. That is, only beings who are (at least in some respects) Quintessentialists would speak a language whose terms behave as Putnam suggests (or at least share Putnam's intuitions that the terms would behave in the relevant ways). If our psychology was relevantly different, then we would not have Twin-Earth intuitions—even if the world does indeed conform metaphysically to the Kripke/Putnam essentialist view of it. I will then review evidence that the world does not in fact conform to Kripke/Putnam essentialism, and argue that the relevant intuitions derive solely from our quintessentialist outlook. As one might put it, Twin-Earth *intuitions* are driven by what's 'in the head' rather than by what's in the world.

2.2. Kripke/Putnam Essentialism and Quintessentialism

In the wake of this compelling thought-experiment, Putnam proposes an account of natural kind terms. He suggests that natural kind terms can be given "ostensive definitions"—that is, one may provide a definition for a natural kind term by ostending an instance of the kind, and indicating that the term applies to anything that is the same as the instance in important respects. For example, we could define the term "water" by pointing to a glass of it and noting that "water" applies to anything that bears the "same liquid relation" to the ostended sample:

> The logic of natural-kind terms like "water" is a complicated matter, but the following is a sketch of an answer. Suppose I point to a glass of water and say "this liquid is water" . . . My 'ostensive definition' of water has the following empirical presupposition: that the body of liquid I am pointing to bears a certain sameness relation (say, *x is the same liquid as y* . . .) to most of the stuff I and other speakers in my linguistic community have on other occasions called "water" (1975, p. 224–5).

And more generally: "One can give [someone seeking to learn a natural kind term] a so-called ostensive definition—'this (liquid) is water'; 'this (animal) is a tiger'; 'this (fruit) is a lemon' " (1975, p. 229). Kripke makes similar remarks about ostension in *Naming and Necessity* (e.g. p. 135). The crucial thought common to Kripke and Putnam is that underlying essential features determine the extension of "same liquid" and "same substance".

Nathan Salmon (1979) convincingly argues that one would be mistaken to suppose that the essentialist metaphysical doctrine is entailed by the above semantic picture. Rather, Salmon argues, the advocates of the semantic picture make essentialist assumptions from the outset. Indeed, Kripke himself explicitly denies that he ever thought to derive conclusions about the metaphysical status of essentialism from his theory of reference (1980, preface). I am in complete agreement with Salmon that the essentialist metaphysics is presupposed by the semantic analysis, rather than entailed by it. (I argue in the final sections of the paper, however, that the detailed presuppositions are ill-founded.) Here, I would note another implicit assumption in Putnam's argument: namely that the communities of speakers that he describes are, like his readers, Quintessentialists, at least to some extent. Absent such an assumption, the argument would fail. (Unlike his metaphysical assumptions, however, I believe that this assumption is completely correct.)

Imagine a community of speakers—let us call them the Phenomenalists—whose psychology is notably different from the Quintessentialists, especially when it comes to beliefs about the natures of certain kinds and individuals. The Phenomenalists staunchly deny that anything of interest lies below the surface—it would never occur to them to consider two individuals to be essentially similar if they differed in their obvious surface properties. The Phenomenalists thus form their categories and concepts on the basis of perceptual appearance and other such readily-accessible qualities. (In this way, the Phenomenalists take Oscar Wilde's remark "only the most superficial people do not judge by appearances" to be words to live by.)

It should be clear that an ostensive definition of the sort Putnam discusses will yield a term with a very different extension for the Phenomenalists. Putnam writes, "Suppose I point to a glass of water and say '*this* is water,' in order to teach someone the word 'water"' (1975, p. 230). If one attempts to teach a Phenomenalist a term in such as way, the Phenomenalist will acquire a term that he will take to have a very different extension than our term "water" does, at least on Putnam's account of how that term functions. For example, it will strike the Phenomenalist as undeniable that XYZ falls in the extension of this term: it looks like it, smells like it, tastes like it, and so on—and what else matters to the Phenomenalist?

Of course, if we are simply imagining an isolated Phenomenalist attempting to learn English, we might just dismiss his reaction as a semantic error, especially if we believe in public languages whose meanings are determined by community wide dispositions to use terms in certain ways. The response to this is obvious: let us imagine an entire community of Phenomenalists, whose numbers and practices are sufficient to determine a public language of their own. Even if one believes that languages are public and are thus not determined by individual psychological dispositions, there can be no denying that the psychological dispositions of the community of speakers determine which public language they are collectively speaking. If we do indeed speak a language whose term "water" does not include XYZ in its extension, this is because we are psychologically very different from the Phenomenalists.

It might be objected that the Phenomenalists' mindset is at odds with science. That is, haven't we *discovered* that water has an essence (namely H_2O), and so wouldn't it be somehow unscientific of the Phenomenalists not to revise the extension of the term after making such a discovery? Even if we grant the (dubious) scientific claim about the essence of water, it should be clear that this objection has no bite. First of all, no speaker is under an obligation to revise her use of a term in face of such scientific discoveries *unless her community already had a (quint)essentialist view of the term's extension*. If a term in a language has its extension determined by manifest appearance, then scientific discoveries are not relevant to determining its extension (unless of course they bear on facts about manifest appearance). Second, Putnam himself is quite explicit that his claims about meaning do not depend on what is known at a time about the relevant underlying essences. Oscar and Twin-Oscar's words have different extensions in 1750; similarly, "*χρυσός*" (*gold*) as used by Archimedes had its extension restricted to substances composed of the element with atomic number 79—even if Archimedes himself (let us suppose) was unable to tell this substance apart from, say, iron pyrite. It is not the scientific discoveries per se that fix the extension of the terms: it is the dispositions of the community of speakers to intend to use their natural kind terms to carve nature at its (quint)essential joints, whatever those may be. Science is here no more than a post-hoc guide as to the determinate semantic consequences of such dispositions.

If we consider a Quintessentialist community instead of a Phenomenalist one, it is evident that the speakers are likely to speak (or at least believe that they speak) a language of the sort that Putnam describes. If one teaches the word "water" to a Quintessentialist via ostention, then she will readily suppose the word's extension is determined by similarities in underlying quintessence. That is, she will be prepared to limit the use of the term "water" to just those quantities of substance whose quintessences are relevantly similar to that of the ostended substance—and she will be happy to admit that there may be cases where a given quantity of stuff looks just like the original, and yet differs in its quintessence enough that the term "water" will not apply to it. Similarly, if she is taught "tiger" in the same way, she will suppose that the term applies to just those individuals whose quintessences are appropriately similar to the demonstrated individual. And importantly, since their quintessentialist view of the world *preceded* modern scientific discoveries, Quintessentialists would have understood these terms in this way throughout their entire history, just as Putnam supposes that we do.

The Quintessentialists are even able to resolve (or at least appear to resolve) a lingering ambiguity that haunts these ostensive definitions (often referred to as "the qua problem;" for example Devitt and Sterelny, 1987; Dupré, 1981, 1993). There are, for example, a host of non-tigers whose quintessences are in many respects similar to the quintessence of our demonstrated tiger: namely, lions, panthers, cougars, etc. There are also sub-groups of tigers whose quintessences differ in some important and systematic respects: Bengal tigers, Sumatran tigers, Siberian tigers. Let us suppose the demonstrated tiger was a Bengal tiger. How does the Quintessentialist determine that Siberian tigers but not lions have quintessences that are "relevantly similar" to the originally labeled animal? The answer is that she employs a helpful assumption: that novel terms name basic-level categories.[13] This assumption fits very well with the quintessentialist outlook: members of basic-level kinds are taken to have maximally similar quintessences, modulo their also having highly distinctive quintessences relative to other kinds. By settling on the basic-level kind, the Quintessentialist thereby adopts a strategy that maximizes the information an application of a natural kind term to an individual conveys about how the individual's quintessence is similar to and different from other individuals' quintessences. As a consequence, the Quintessentialists favor using this strategy when learning terms via ostension.

Since the Quintessentialists are so accustomed to (unconsciously) applying this strategy, they sometimes leave it wholly implicit in their philosophical works, as they assume that their readers will be Quintessentialists themselves. Put another way, the qua problem can seem not to be that deep precisely because we are Quintessentialists, who privilege basic-level kinds.

If we are indeed Quintessentialists, it is easy to see why our intuitions concerning Twin-Earth and the like accord with Putnam's. Were we, say, Phenomenalists instead, the thought-experiment and its conclusions would gain no traction with us. It is thus at least a necessary condition for Putnam's account to be successful that we think like Quintessentialists in the relevant respects. Is it also sufficient, or does the world have to cooperate? Do we have Twin-Earth intuitions even when the kind in question has no essence? More generally: might we have the *intuition* that members of a particular kind *must* share an essence even if in fact they do not? If so, perhaps not only are the specifically Twin-Earth intuitions driven by quintessentialism, but so too is the bedrock idea that the kinds that correspond to our natural kind terms group individuals by their essences.

2.2.1. The Tragic Mulatto

Before considering whether the kinds that Putnam and Kripke discuss actually do group their members by essence, it is worth recalling a nineteenth and twentieth-century North American literary and cinematic trope, the Tragic

[13] A separate question is how a Quintessentialist figures out which categories are basic level. There is a wealth of interesting work on this issues (e.g. Murphy, 2002; Rosch, 1973, 1975, 1978; Rosch, Mervis, Gray, Johnson, and Boyes-Braem, 1976). Shape is an important (though defeasible) guide. It is important to separate this empirical question from the point that a Quintessentialist, who can *somehow* distinguish basic-level from non-basic-level kinds, even when the kinds in question are novel, can use this ability to fix on a meaning for a term introduced via Kripke/Putnam style ostension.

Mulatto (or Tragic Mulatta, as the trope is sometimes styled). The trope typically involves a young woman who "looks and acts like a white person," and in many cases believes herself to be white. Everyone around her unquestioningly believes that she is white, and she rises to an enviable position in society and is sought after by eligible young white men. Then the 'terrible' news is discovered: the woman is not, in fact, white, but rather has mixed racial ancestry. The discovery of one black ancestor spells her ruination. As the literary trope has it, a 'single drop' of 'black blood' suffices to make one black; she loses everything overnight, is rejected by friends and lovers, and cast out of white society. In some instances, the tale ends with her being sold into slavery.

This depressing trope was very popular, and clearly was not in any way difficult for its readers to comprehend. Let us consider its structure then: it requires the reader to suppose than a person can appear in every observable respect to be white, and yet in fact not be white, but rather black. This is something that would be incomprehensible to the Phenomenalist about race, yet is immediately comprehensible to the racial Quintessentialist. Since quintessence only defeasibly causes its bearers to have their characteristic observable properties, the possibility exists for a member of a kind to have the relevant quintessence, and yet not share any of the characteristic properties of the kind. Conversely, an individual may have the characteristic superficial qualities associated with a kind, and yet in fact have the quintessence of another kind. (Compare Gelman and Markman's (1986) study, where preschoolers readily accepted that an individual could be a member of chemical or biological kind, despite better resembling members of a different kind.) It is this quintessentialist understanding of race that the Tragic Mulatto trope exploits.

The Tragic Mulatto trope has much in common with the Twin-Earth thought experiment; someone whose mindset allowed them to accept the Tragic Mulatto tale hook, line, and sinker would surely also have Twin-Earth intuitions about race. If there can be one "tragic mulatto," why not a whole planet? Why not a Twin-Earth populated by individuals who look and act "like white people," but are in fact, like the Tragic Mulatto, *really black*? Just as the Tragic Mulatto's suitors incorrectly apply the term "white" to her, a visitor to this Twin-Earth would incorrectly apply the term "white" to its inhabitants, and so on, so forth. The same conceptual structure allows for both the Twin-Earth example and the Tragic Mulatto trope. Both hinge on the idea that something can appear in every respect to belong to a kind, but in fact fail to, since the individual lacks the relevant essence.

There is, of course, no such thing as black or white essence. Racial groups are social constructs: they are not based on any biologically real essences. Genetic variation within racial groups is just as high as it is across racial groups—there is no such thing as "black DNA," and there is most certainly nothing intrinsic that would set apart a young woman as *truly* being black, despite having pale skin and other phenotypic features associated with white people.[14] In this case the world does not vindicate the belief in hidden essence, but racial groups have historically been—and continue to be—highly quintessentialized. These quintessentialist beliefs suffice for the analog of the Twin-Earth intuitions in the case of race. This suggests that quintessentialism may be not only necessary, but also sufficient for the relevant intuitions.

2.3. Do Biological Kinds have Essences?

One outstanding question at this point is whether the kinds that Putnam and Kripke actually discuss turn out to have essences, or whether—as in the race case—both our Twin-Earth style intuitions and our confident belief in kind essence are simply driven by unjustified quintessentialist beliefs. For a kind to "have an essence" in the Kripke/Putnam sense, there must be hidden underlying features that are necessary and sufficient for kind membership—features of the sort that science investigates and discovers. Moreover, for a natural kind term to have the sort of extension that Kripke and Putnam suppose, it is necessary that any two members of the relevant kind (including any two quantities of the relevant substance) be similar in some such scientifically discoverable and essential respect. Putnam is very explicit that *having chemical composition* H_2O is the relevant essential respect in the case of *water*, and also that *having the appropriate genetic code* is the relevant essential respect in the case of *lemons*. He writes:

> [A] critic has maintained that the *predominant* sense of, say, "lemon" is the one in which anything with... the superficial characteristics of a lemon is a lemon. The same critic has suggested that having the hidden structure—the genetic code—of a lemon is necessary to being a lemon only when "lemon" is used as a term of science. Both of these contentions seem to me to rest on a misunderstanding...
>
> The sense in which literally *anything* with the superficial characteristics of a lemon is necessarily a lemon... is extremely deviant... At the same time the sense in which to be a lemon something has to have the genetic code of a lemon is *not* the same

[14] Some researchers have recently claimed that race does, in fact, have a biological reality, in particular that racial divisions can be at least approximately cashed out in terms of continent-based ancestral breeding populations, which affect to some extent the nature and frequencies of some alleles (see, e.g. Kitcher (2007) for sympathetic discussion). Such findings, even generously interpreted, do not amount to "racial essences" or "b lack DNA" or the like, as Kitcher makes very clear. For example, the claims in question are statistical/probabilistic in nature, and quite weak ones at that: only .0005% of human genetic variation is even putatively explained by membership in these groups (Maglo, 2011). It is also far from clear that these findings hold up to scrutiny (as opposed, perhaps, to simply being expressions of quintessentialism); there are a number of reasons to be skeptical regarding them—reasons which range from concerns about the statistical methods employed to objections to the theoretical interpretations bestowed (for an excellent review, see Maglo (2011), and sources cited therein). Further, the Tragic Mulatto is specifically characterized by having largely white ancestry but a "single drop of black blood," which suffices to make her black. Nothing in these statistically based genomic analyses would deliver this result (in fact one criticism of them is that they tend to abstract away from cases of mixed ancestry, thus potentially making the findings appear more clean-cut than they in fact are).

as the technical sense (if there is one, which I doubt). The technical sense, I take it, would be one in which "lemon" was *synonymous* with a description which *specified* the genetic code. But when we said ... that to be *water* something has to be H_2O we did not mean ... that the *speaker* has to *know* this ... similarly, even though the predominant sense of "lemon" is one in which to be a lemon something has to have the genetic code of a lemon (I believe), it does not follow that "lemon" is synonymous with a description which specifies the genetic code explicitly or otherwise. (1975, pp. 239–240, original emphasis).

Kripke makes similar remarks concerning the "internal structure" of tigers (1980, pp.120–1), and later speaks of "scientific discoveries of species essence" (p. 138). Does the world cooperate with these claims, or are these quintessentialist misinterpretations of science? That is, are Kripke and Putnam simply reporting scientific facts in their discussions, or are the discussions fueled instead by unfounded quintessentialist convictions that natural kinds *must* have essences? Might the very claim that natural kinds have scientifically discovered essences be no more than an articulation of inchoate quintessentialist intuition?

2.3.1. Species and Sex; Genes and Essence

On the question of whether species have these sorts of essences, there is a degree of consensus among philosophers of biology (and indeed biologists) that is almost unprecedented in philosophy at large (e.g. Dupré, 1981, 1993; Ghiselin, 1987; Hull, 1965; Laporte, 1997, 2004; Mayr, 1982, 1988, 1991; Okasha, 2002; Sterelny and Griffiths, 1999, and many others.). There is no such thing as "lemon DNA," no common genetic code that makes for membership in the kind *Panthera tigris*.[15]

To a first approximation, most biologists subscribe to the notion that a species is delineated by the boundaries of an ecological niche, or by the boundaries of a reproductive community.[16] Within such bounds, considerable genetic variation is possible, and conversely, it is possible for there to be *less* genetic variation *across* such boundaries. As a concrete illustration, consider the cutthroat trout (*Salmo clarki*). This species has a number of subspecies, and there is considerable genetic divergence between these subspecies. In a study of seven subspecies of cutthroat trout, zoologists Fred Allendorf and Robb Leary report that:

A highly variable pattern of genetic divergence exists among the seven subspecies ... Very little genetic divergence exists among Colorado, Snake River, greenback, and Yellowstone cutthroat trout. Nei's genetic distances between these subspecies are

[15] In what follows, I will limit my discussion to the species-level in the taxonomy, since this is, I think, in keeping with Kripke's and Putnam's intentions.

[16] These distinct ways of characterizing species are considered in more detail in the section headed "*Extrinsic Species-Essence and Plenitude*".

typical of those reported for conspecific populations in a diversity of freshwater and anadromous fishes (Avise 1974; Avise & Smith 1977; Buth & Burr 1978; Loudenslager & Gall 1980; Buth et al. 1984). In contrast, substantial biochemical genetic divergence exists between coastal, Lahontan, and westslope cutthroat trout and between these fishes and the other four subspecies. These genetic distances are truly exceptional for conspecific populations; *they are similar to or larger than values observed between many species of fish* (Johnson 1975; Avise & Ayala 1976; Buth & Burr 1978; Phelps & Allendorf 1983; Yates, Lewis, & Hatch 1984). *Surprisingly, the coastal, Lahontan, and westslope cutthroat trout are as similar or more similar to the rainbow trout than they are to the other subspecies.* (Allendorf and Leary, 1988, pp. 172–3, emphasis added)

Thus a member of one species (e.g. a member of the Lahontan subspecies) may have more genetically in common with a member of another species (namely rainbow trout) than with a member of its own species (e.g. a member of the Snake River subspecies). While such findings may not be the norm in biology, this is not a wholly isolated occurrence; for example, high degrees of intra-species genetic divergence can be found among various other Salmonids as well (Pennell and Barton, 1996). As we shall see, these are not just weird cases: biology systematically confounds our quintessentialist convictions.

As against this, recently Michael Devitt (2008) has criticized the consensus in the philosophy of biology, and indeed in biology and zoology, and argued for the existence of intrinsic, microstructral species essences of the Kripke/Putnam variety. It is instructive to consider Devitt's arguments, since he articulates a view that is, I think, widely held in philosophy (outside of philosophy of biology). He argues that biologists and zoologists are *committed* to positing such essences—even if they assert otherwise—on pain of explanatory inadequacy. Devitt writes:

There has to be something about the very nature of the group—a group that appears to be a species or taxon of some other sort—that, given its environment, determines the truth of the generalization [e.g. that *Indian rhinos have one horn* and *African rhinos have two horns*]. That something is an intrinsic underlying, probably largely genetic, property that is part of the essence of the group. Indeed, what else could it be? Some intrinsic underlying property of each Indian rhino causes it, in its environment, to grow just one horn. A different such property of each African rhino causes it, in its environment, to grow two horns. The intrinsic difference explains the physiological difference. If we put together each intrinsic underlying property that similarly explains a similar generalization about a species, then we have the intrinsic part of its essence. (2008, p. 355)

And summarizing:

I have presented a positive argument for Intrinsic Biological Essentialism. We might sum it up: *structural explanations in biology demand that kinds have essential intrinsic properties.* (2008, p. 365, original emphasis)

Devitt's argument is intuitively appealing, but not, I think, ultimately successful (see Barker (2010) and Ereshefsky (2010) for further discussion). First of all, the argument is undermined by the fact that members of the same species often do not share the same phenotypic traits; conspecificity is compatible with a great deal of variation in phenotype at a time, and even more dramatically over time. That is, it is possible for a species to radically change its phenotypic features over time without a speciation event occurring. In the extreme case, members of a species at t1 might not have any non-trivial phenotypic properties in common with the members of the species at t2. However, in what follows I will set aside this particular concern, and consider only cases in which a particular phenotypic trait actually is shared by members of a species.

Devitt is, of course, indisputably correct that *each particular African rhino* has some intrinsic features that, in combination with the environment, are causally responsible for that individual's having horns. This does not entail, however, that those very same intrinsic features are also responsible for other African rhinos' having horns. Whether this is so is a substantive empirical hypothesis, not one whose truth can be intuited in advance. For a simple illustration of the general point, let us revert to a well-known chemical case, that of jade. There are two very different chemical compounds that both fall under our term 'jade': jadeite ($NaAl(SiO_3)_2$) and nephrite ($Ca_2(MgFe)_5Si_8O_{22}(OH)_2$). Substances composed of these two respective chemical compounds have very similar observable properties (in fact they are far more similar than many conspecific plants and animals). Thus a sample of jadeite and a sample of nephrite will have many properties in common (in fact they will be indistinguishable to all but the most experienced artisans). It is also true that the observable properties of the sample of jadeite are determined by its intrinsic chemical structure in conjunction with the environment, and similarly for the sample of nephrite. Yet there will be no common intrinsic chemical structure that explains the shared features of the two samples of jade.

(There will, of course, be the non-explanatory disjunctive property of being composed of *either* $NaAl(SiO_3)_2$ *or* $Ca_2(MgFe)_5Si_8O_{22}(OH)_2$.However, it is important to see that disjunctive properties cannot play the explanatory role that Devitt has in mind, or else the whole enterprise is trivialized. For example, let us suppose with Devitt that there is a common intrinsic property had by tigers that explains why they are striped. Let us also suppose that there is a different common property that explains why canna lilies are striped. If disjunctive properties are allowed to figure as common intrinsic explanatory properties in Devitt's sense, then there will be a further shared intrinsic property that explains why *this tiger* and *this lily* both have stripes. If disjunctive properties are countenanced in this endeavor, then *shared* properties become far too cheap to be of interest. Certainly, it would not then be a *biological* hypothesis that a common property explains why Indian rhinos have one horn—it would simply be a familiar point about the logic of disjunction.)

In the case of jade, *we* have selected these two chemically different but manifestly similar substances to figure in some of our statues and jewelry; but natural selection can do something similar when it comes to phenotypic features. Consider, for example, the property of having typical external human female genitalia—the property that is overwhelmingly used to classify humans, particularly newborn infants, as female.[17] What is the genetic basis for having such a property—what intrinsic genetic basis explains the possession of such a property (as Devitt might put it)? This is a question that many educated adults believe they know how to answer: it is having two X chromosomes, they say! Possession of two X chromosomes (or more accurately, possession of at least one X chromosome and no Y chromosome), it is thought, is both necessary and sufficient for being a human female, and this further explains the possession of typical female characteristics. The distinctive essential feature is thus not only thought to be necessary and sufficient for having typical external female genitalia, but it is also explanatory in the sense that Devitt seeks.

This proposal does not in fact withstand scrutiny. While the majority of females do not have Y chromosomes, there are some that do; a small but substantial portion of the population have 46, XY karyotypes, but have typical external female genitalia.[18] There are numerous ways in which this situation can arise. For example, various mutations can lead to *androgen insensitivity*, meaning that that cell receptors do not respond to androgenic hormones, so cells are unaffected by the presence of these hormones. Alternatively, there can be any number of changes to the biochemical processes involved in producing these androgens: for example, there may be a range of alterations affecting the synthesis of testosterone from cholesterol, or there may be changes in enzyme levels that prevent testosterone from being broken down into dihydratestosterone. For each of the determinate ways in which this developmental pathway can be altered, there are a range of different genetic mutations, or combinations of mutations, that may be responsible. Further, this is far from an exhaustive list of the ways in which a 46, XY karyotype may be associated with typical external female genitalia (Kolon, 2008).[19]

[17] This discussion in this section is concerned with the biological categories of sex (e.g. male/female), not with the social categories of gender (e.g. man/woman). The two are, of course, often confused, in accordance with our tendency to quintessentialize gender to a very high degree. While it is often pointed out that the latter has no intrinsic essential bases, the corresponding point is less frequently made in the case of the former. The discussion here should illustrate that even the seemingly simple categories of sex are in fact highly complex and resistant to delineation in any simple essentialist terms.

[18] This discussion abstracts away from cases of individuals with ambiguous genitalia, which can (though need not) result from many of the conditions described here, and from other conditions besides.

[19] As a further level of complication, consider the phenomena of mosaicism and chimerism. These are conditions in which the cells within a given individual have different genetic makeups. For example an individual with mosaicism/chimerism might have some cells with karyotype 46, XX and others with 46, XY, or some with 47, XXY and others with 46, XY, or some with SRY mutations and others without, and so on. The effect on the phenotype depends on a range of factors, for example the relative proportion of the different cells, their location and degree of distribution around the body, and so on so forth. For example, if someone has a mix of 46, XX and 46, XY cells (which is rare but possible), they may have typical male genitalia, typical female genitalia, or ambiguous genitalia, depending among other factors on the proportion and location of the two cell types.

To return to Devitt's argument, suppose we have a number of newborn females, all with typical external genitalia, and suppose we seek to explain why their genitals are formed as they are. For each individual infant, there will be a genetic component to the explanation (as well as an environmental explanation), but this genetic component need not be the same for each infant. There may be considerable overlap in explanation (the majority of individuals with female genitalia have 46, XX karyotypes, and—crucially—lack the factors that lead such a karyotype to be associated with typical male genitalia), but in the case of some individuals, the explanation will go by a rather different route. One cannot infer a universal and essential intrinsic genetic component on the basis of shared morphological characteristics, even within a given sex of a given species.

Of course, Quintessentialists that we are, we find such inferences appealing to the point of being almost irresistible. As a result, one might be tempted to rescue the idea by 'setting aside' cases of 46, XY females, so as to preserve the idea that there is a genetic essence shared by all and only females that explains their typical morphological features. (I want to be absolutely clear that I am in no way attributing such a response to Michael Devitt, however.) There is no question that the individuals in question have the relevant morphological feature—typical external female genitalia—so if the explanation of why *females* have this feature is to make reference to a genetic essence shared by all females, the only option is to insist that these individuals are not, in fact, female. That is, if we insist that only 46, XX individuals are female, then we can rescue the idea that the kind *human female* has an essence, and further that this essence explains the typical morphological features of the category. The *real* females are the 46, XX females, the others are aberrations, freaks, not really female. I say this is no more than quintessentialism in its most unjustified and pernicious form. Biology does not provide neat, essentialized categories of this sort, however much quintessentialism demands them.[20] There are a number of different biological routes to the phenotype, and the natural world does not label one or another of these as the genuine, true, real, or normatively right one.[21] Sadly, a number of female athletes have had their careers and even their lives ruined by the quintessentialist way of thinking, which is blind to this natural fact.

Further, the sort of phenomenon illustrated here is widespread in the biological world. In general, within a given species, individuals who share a common phenotypic feature need not share intrinsic microstructural bases that gave rise to the feature. One simple illustration of this phenomenon

[20] For a number of years, 46, XY females were referred to in the clinical literature as "male pseudohermaphrodites," a label which is no longer in use. Even in the context of medicine, quintessentialism can leave its traces, but it is hard to maintain in the face of a real investigation of nature.

[21] Ah, but the variations described in this section tend to be incompatible with reproduction—is this not biology's way of signaling that these individuals are not really female? Not more so here than with any genetically based condition that results in infertility.

involves what are known as *phenocopies*—individuals who display environmentally induced, non-hereditary phenotypic features which are identical to genetically induced, hereditary features displayed by other individuals. This can occur within a particular species—for example, the Common Rabbit (*Oryctolagus cuniculus)* frequently sports black fur. This trait—black fur—can arise relatively straightforwardly from a given rabbit's genetic make-up—that is, as an inherited trait that manifests itself across various environments. The trait can also occur, however, among the Himalayan breed under certain environmental conditions. Himalayan rabbits, when raised in moderate temperatures, have white body fur with black tails, noses, and ears; if they are raised in cold temperatures, however, they develop wholly black fur. The genetic, developmental and biochemical pathways to black fur differ significantly between the two cases (Baum et al., 2010; Sawin, 1932); there is no common intrinsic explanation for the possession of black fur by the members of *O. cuniculus*. (And conversely, there need be no genetic difference between a Himalayan rabbit with black fur and one with mostly white fur.)

Perhaps one might be tempted to think that this example is misplaced because *having black fur* is not a species-wide characteristic of rabbits, or perhaps because appeal was made to a particular breed of rabbit (which, one might be tempted to think, *would* be such that its members share an intrinsic microstructural basis for the trait). This is not so. Consider, for example, that *having three toes on the hind feet* is a characteristic property of guinea pigs (*Cavia porcellus*). Possession of this phenotypic property is due to a flexible interaction-effect between a number of factors, both genetic and non-genetic—that is, there is a generous range of pathways, all of which lead to having three hind toes. Individual guinea pigs can differ significantly from each other with respect to these factors and yet each have three toes, as was demonstrated in a series of classic experiments conducted by Sewall Wright (1934).

Alternatively, consider the salamander *Ambystoma talpoideum*. *A. talpoideum* can undergo either metamorphosis from its larval stage to its adult stage, or else can undergo pedomorphosis, that is, it becomes sexually mature while retaining its larval characteristics. A general characteristic property of *A. talpoideum* is that it tends to undergo pedomorphosis if it is raised in a favorable aquatic environment. However, the genetic basis for this property has been found to differ between two populations of *A. talpoideum* that are separated by as little as 15 km (Harris, Semtlisch, Wilbur, and Fauth, 1990). The developmental pathway leading to metamorphosis can be disrupted in any number of ways (e.g. by increased prolactin secretion, by blockage of thyrotropin-releasing hormone secretion, by reduced sensitivity to thyroxine, or by other means), so the differences in the genetic basis of pedomorphism likely translate into significant differences in the biochemical processes undergone by individuals from the two populations (Harris et al., 1990; West-Eberhard, 2003). Further, the two populations of salamanders are such that the females typically produce approximately eighteen eggs each; however, once more the genetic basis for this property differs between the two populations (Harris et al., 1990).

Such examples abound. Phenotypic traits are the upshot of complex biochemical processes controlled in most cases by a number of genes. Differences in the genetic level need not translate into differences in the biochemical processes, and even differences in the biochemical processes need not translate into differences in phenotypic traits (see, e.g. Zuckerkandl and Villet (1998) for detailed discussion of the latter point; for example, in some cases there may be an important trade-off between chemical affinity and chemical concentration).

More generally, biologists and geneticists often make use of the notion of *canalization*—the idea that the development of a phenotypic trait, particularly one that is important to survival, can arise despite variation in genes, environment, and the resulting developmental pathways. Canalization of a trait insures that the trait is stably expressed in the face of underlying genetic variation (e.g. Gibson and Wagner, 2000; Waddington, 1942).[22] To put the point in what is perhaps more familiar terminology, we might say that phenotypic traits often exploit a certain *multiple realizability* at the microstructural level.

Notice that it would not be an adequate defense of Devitt's idea to say that the human female, the trout, the rabbit, the guinea pig, and the salamander may be the less common sort of case, and that phenocopying and canalization may not be universal phenomena. Whether or not this is true, it is not the sheer number of such cases, but the way in which they subvert the idea that species as such *must* meet the relevant (quint?)essentialist standards for being a respectable natural kind.

The general theoretical point is this: Every macroscopic phenotypic property depends on a massive number of biochemical reactions, originating with the genes themselves but continuing along the entire developmental pathway, at each point potentially subject to environmental influences, influences from other genes, and so on. If there was a one-to-one correspondence between sameness of macroscopic phenotypic properties and sameness of biochemical pathway, then—simply put—*there would be considerably more phenotypic variation than there in fact is.* The idea that one may generally infer shared 'underlying' features from a shared phenotypic feature just does not pass empirical muster. Nature is, fortunately, far more robust.

2.3.2. Extrinsic Species—Essence and Plenitude

If species do not have intrinsic essences, if there are no general microstructural genetic facts of the sort that Devitt, Putnam and (more offhandedly) Kripke posit as making for species membership then does this mean that species have no essences whatsoever? Recently, several philosophers of biology have argued that species can indeed be said to have essences, albeit extrinsic ones.

[22] Consider, for example, the wide range of developmental pathways that can all lead to typical external female genitalia, as discussed.

For example, Samir Okasha (2002) and Joseph LaPorte (1997, 2004) and others have argued that species do indeed have essences, in the sense that there are necessary and sufficient conditions for belonging to a given species. These necessary and sufficient conditions are relational, however, and do not yield the result that species-membership is an essential property of the individual (e.g. it will not be an essential or even necessary property of Socrates that he be a member of *Homo sapiens*, contra Wiggins (1980)). Nonetheless, if such species essences are to be found, then there would at least be a determinate *same animal* relation, of the sort that Kripke and Putnam appeal to. (This result would be arrived at only on the basis of some charitable reconstruction—in particular we would need to supplement the discussion of ostensive definition with the notion that the *same animal* relation in fact amounts to the *same species* relation, which is in effect to grant a solution to the qua problem (Devitt and Sterelny, 1987). Since basic-level kinds do not map neatly onto species, or in some cases onto any taxon (Dupré, 1993), this is already quite a generous amendment to the view.)

Broadly speaking, on most of the popular accounts, to be a member of a given species, for example *Panthera tigris*, is to be part of a particular chunk of the genealogical nexus, which begins with a particular speciation event, and ends with either another speciation event, or with extinction. A more specific version of the species concept is the phylogenetic (or cladistic) species concept. LaPorte offers the following as an elaboration of the phylogenetic species concept: "a species name like '*Panthera tigris*' is to be defined something like as follows: '*Panthera tigris* $=_{\text{df}}$ the lineage descending from ancestral population P and terminating in speciation or extinction,' P being . . . an appropriate population in the lineage that gave rise to today's tigers." (2004, p. 54)

Of course we are still owed an account of what a speciation event consists in, as many philosophers of biology have noted (e.g. Dupré, 1993; Kitcher, 1984; LaPorte, 2004; Okasha, 2002; Sterelny and Griffiths, 1999). There are two dominant lines of thought here: that speciation events occur when populations become reproductively isolated from each other, and that such events occur when populations come to occupy distinct ecological niches. Both lines of thought are frequently appealed to by biologists and zoologists, and each plays an important theoretical and explanatory role in at least some contexts of biological inquiry.

Before considering the significance of having multiple scientifically useful ways of characterizing speciation events, it is worth briefly noting that, whatever details one embraces, species membership is not an intrinsic matter (see, e.g., Dupré, 1981, 1993; Kitcher, 1984; LaPorte, 1997, 2004; Okasha, 2002; Sober, 1980; Sterelny and Griffiths, 1999). For example, the dominant accounts of the species concepts deliver the result that, if two populations are geographically separated in such a way as to prevent interbreeding, and result in the two populations occupying distinct ecological niches, then these populations constitute distinct species (at least on many tenable construals of these speciation conditions). No reference is here made to the intrinsic properties of the members of the species; thus the possibility remains that two individuals with identical intrinsic properties could belong to distinct species (as a result of belonging to such geographically separated populations). As another (somewhat simplified) illustration, consider the case of an individual with significant genetic mutations that render the individual incapable of either interbreeding with his parents' population or occupying their ecological niche. If there are other similar individuals around to form a distinct interbreeding population or co-occupy a distinct ecological niche, then a speciation event may be underway, and the individual—sometimes termed a "hopeful monster"—may count among the first population of a new species. If no such fellows are to be found, however, the individual will simply count as an atypical and infertile member of his parents' species.[23] Thus, even if we grant that we may be able to specify necessary and sufficient conditions for membership in a species, we are very far indeed from the sort of essentialism Kripke and Putnam had in mind (LaPorte, 2004; Okasha, 2002).

By now it should occasion no surprise to note that the pursuit of extrinsic essential conditions for species membership quickly runs into its own complications. As the above discussion indicates, the phylogenetic species concept can be supplemented in at least two useful and explanatory ways—that is, one can characterize speciation events as arising when two populations cease to interbreed, or when two populations come to occupy distinct ecological roles. These two characterizations of speciation events can lead to competing results: for example, it is possible for two populations to occupy different ecological niches and yet still interbreed, and so on. Further, the phylogenetic species concept (with its premium on shared ancestry) is not even the only concept on the table; there is, for example, the biological species concept (BSC), which takes interbreeding to be the sole determiner of species membership. These two different species concepts may deliver different results, even if the phylogenetic species concept is supplemented by the interbreeding approach to the speciation question. As an illustration, suppose a species splits into two (i.e. non-interbreeding) species S1 and S2 at time t1. Then later at time t2, S1 and S2 cease to be reproductively isolated from each other, and form an interbreeding population. The phylogenetic species concept will still entail that there are two distinct species at t2 (distinguished by their different ancestry going back to time t1), whereas the BSC in its most straightforward form would posit a single species at t2. A further species concept takes occupancy of an ecological niche to be the sole determiner of species membership, rather than simply a determiner of speciation, and so on so forth.[24]

[23] Devitt (2008) writes as though the standard species concepts are not in conflict with his Intrinsic Biological Essentialism. The fact that the standard species concepts allow for such cases suggests otherwise, however (Barker, 2010; Ereshefsky, 2010). Thus Devitt's proposal is highly revisionary.

[24] Devitt (2008) correctly points out that these species concepts fall short of telling us what it is to be a *tiger*—that is, while they may tell us what it is for a *population* to be a *species*, or what it is for *two organisms* to be *conspecific*, they do not tell us what it is for an *organism*

Notably, despite the fact that these competing concepts deliver different results about whether two individuals are conspecific or not, and also about how many species should be recognized, and so on, they all have a claim to being scientifically useful. This has led a number of prominent philosophers of biology, beginning with Philip Kitcher (1984) and John Dupré (1981, 1993) to argue for pluralism concerning species concepts (see also Ereshefsky, 1998; Holter, 2009; Rosenberg, 1994; Stanford, 1995, and others).[25] The extent of the pluralism is not limited to the range of species concepts laid out above either; for each of the seemingly particular species concepts, there are any number of ways of spelling out the details. Kitcher notes this in a recent paper:

> I proposed [in 1984] that there were many different species concepts, appropriate for different purposes of inquiry Both Dupré and I, however, tended to think in terms of manageable pluralism, or limited promiscuity; for my part, I took the Biological Species Concept to be one among a number of contenders. The real trouble, however, is that the Biological Species Concept itself allows for indefinitely many ways of development, depending on how one approaches the notions of population and of reproductive isolation. (2007, pp. 300–301)

The arguments of Kitcher and his colleagues on this point seem to me completely convincing: if we seek to characterize species—to give necessary and sufficient conditions for membership in a species—then we will need to countenance a plenitude of sets of such conditions, each of which is scientifically useful and appropriate. Thus even the most charitable attempts to reconstruct essentialism about species face the problem of there being *too many* candidate

to be a *tiger* (or a lion or a zebra, and so on). Some qualification is in order here: we cannot produce any such *purely qualitative* specification of the essence of a kind like the tiger or the lion. However, if we are allowed to directly refer to particular individuals—e.g. a particular founding population—then we can provide such necessary and sufficient conditions (e.g. to be a tiger is to be descended from *this* ancestral population prior to any further speciation events occurring among the population's descendents).

Devitt's point here is correct, though I disagree with his interpretation to the effect that this indicates a significant hole in biology/philosophy of biology. An alternative interpretation would simply be that, in light of the best science and philosophy, we have found that full-blooded, purely qualitative necessary and sufficient conditions are only to be found at the level of *what it is to be a species*, or *what it is to be conspecific*, not at the level of *what it is to be a tiger*. Of course, since tigers constitute a species, this places some necessary conditions on being a tiger, though they will fall short of being also sufficient conditions.

This state of affairs parallels a relatively common view in the discussion of philosophical essentialism: many philosophers hold that we can provide qualitative necessary and sufficient conditions for, e.g. *being a person* (where this may be distinct from being a member of the species *Homo sapiens*; compare the literature on personal identity), and further that Socrates is essentially a person. However, they do not hold that there are qualitative necessary and sufficient conditions for *being Socrates* (though there may be non-qualitative ones, e.g. *coming from* this *sperm and* this *egg*). The situation in philosophy of biology does not seem worryingly different.

[25] There are a number of points of disagreement and discussion in the literature, for example whether pluralism is compatible with realism about species, and I will not attempt to reconstruct all the subtleties here.

species essences for us to be able to say: *here* is the essence of the kind *Panthera tigris*—the set of properties that all and only actual and possible members of the kind possess. Biological science has not uncovered such hidden conditions that properly govern our use of the term "tiger;" biology does not issue in a privileged *same animal* relation. At best it has uncovered a multitude of distinct sets of conditions and relations, each with equal claim.[26]

2.4. Chemical Kinds and Essence

Kripke/Putnam essentialism about biological kinds looks to be untenable, even if one attempts to reconstruct it in the most charitable of ways. But what of chemical kinds, such as water and gold? Surely *here* the world cooperates with our quintessentialist intuitions. After all, don't we *know* the essences of water and gold to be H_2O and the element with atomic number 79 respectively? More generally, has not chemistry actually discovered a privileged *same substance* relation (namely being either the *same element* or the *same compound*)—one which could be pressed into duty in the following sort of manner:

> Suppose I point to a glass of water and say 'this liquid is called water' ... My "ostensive definition" has the following empirical presupposition: that the body of liquid I am pointing to bears a certain sameness relation (say, *x is the same liquid as y*, or *x is the same*$_L$ *as y*) to most of the stuff I and other speakers in my linguistic community have on other occasions called 'water' ... The key point is that the relation same$_L$ is a *theoretical* relation: whether something is or is not the same liquid as *this* may take an indeterminate amount of scientific investigation to determine. Moreover, even if a "definite" answer has been obtained either through scientific investigation or through the application of some "common sense" test, the answer is *defeasible*: future investigation might reverse even the most 'certain' example. (Putnam, 1975, pp. 224–5)[27]

[26] I believe that this is anyway a widespread consequence of philosophical essentialism in its most general form—i.e. the thesis that there are two ways in which an item may have a property, namely essentially or accidentally. Put very crudely, where one might suppose there is a single item with *n* properties, essentialism has difficulty avoiding the consequence that there is something on the order of the *cardinality of the power of set of n* items instead, differing only in which of the *n* properties are had essentially vs accidentally. While a number of refinements are in order, essentialism without plenitude on a vast scale is, I think, an untenable position. For the details and the arguments, see Leslie (2011, submitted).

[27] Here and elsewhere, Putnam speaks of a "same liquid" relation. The appeal to chemistry is more plausible, however, if this is characterized as a "same substance" relation instead. For example, the dominant use of the term "water" in chemistry is a phase-neutral one, so that water can also occur in solid and gas phases. Nothing in the criticisms of essentialism that follow hang on this adjustment; if anything, this adjustment makes the case for essentialism more plausible. For example, the oft-made claim that "water is identical to H_2O" is clearly false as stated if "water" is a phase-specific term. For then, by parity of reasoning and the transitivity of identity, the absurdity that water is identical to ice is derivable. See Johnston (1997). In what follows, I will use the term "water" in a phase-neutral sense, unless otherwise specified.

The idea is elaborated in the following manner:

[A] liquid with the superficial properties of "water" but a different microstructure *isn't really water* . . . Suppose, now, that I discover the microstructure of water—that water is H_2O. At this point I will be able to say that the stuff on Twin Earth that I earlier *mistook* for water isn't really water. (Putnam, 1975, pp. 232–3)

As this passage makes clear, Putnam believes that the *same substance* relation is determined by 'microstructural' properties, and further that 'H_2O' specifies such a microstructure. Also emphasized in these last two passages is the notion that these microstructural properties—these determiners of the *same substance* relation—are *discovered*. The same theme is also found in Kripke, who makes it clear that we may not only discover, contrary to what we believed, that two substances are distinct, but also that two substances are in fact one and the same:

[I]f this substance [H_2O] can take another form—such as the polywater allegedly discovered in the Soviet Union with very different identifying marks from that of what we now call water—it is a form of water because it is the same substance, even though it doesn't have the appearance by which we originally identified water. (1980, pp. 128–9).

And more generally:

[S]uppose some items (let the set of them be *I*) are discovered and are believed to belong to a new kind *K*. Suppose later it is *discovered* that the items in *I* are indeed of a single kind; however, they belong to a previously known kind, *L*. Observational error led to the *false initial belief* that the items in *I* possessed some characteristic *C* excluding them from *L*. (1980, p. 136, emphasis added).

The basic idea behind this essentialist model for chemical kinds is clear: there is a privileged relation *same substance as* that science discovers, and which determines the extension of our natural kind terms. Further, this scientific discovery has already been made, at least in the case of water and gold, which have been found to be identical to H_2O and the element with atomic number 79 respectively. It may now seem reasonable to infer from these examples and others that, for chemical kinds, the *same substance* relation can be characterized as a) same atomic number in the case of elements, b) same chemical formula in the case of compounds. Such 'microstructural' essences are the determiners of the extension of chemical kind terms, and macroscopic manifest features are at best a guide to these true essential features.

We can now pose the question: does science *actually* deliver such a privileged *same substance* relation, or do we simply have once again the quintessentialist intuition that science *must* do so? That is, does the Kripke/Putnam discussion simply retell the scientific facts, or is the science once again distorted, perhaps by quintessentialist convictions?

In this section I will argue that the relevant sciences deliver no such privileged *same substance* relations. The idea that same atomic number/same chemical formula captures this relation simply will not do, especially in the case of compounds; nor is there a unique privileged substitute relation. At best, science delivers a number of candidate *same substance* relations, *many of which rely on the macroscopic and the manifest rather than the microstructural.* Finally, I discuss some examples that suggest that the chemical kinds of thought and talk may not line up especially neatly with the kinds of chemistry, suggesting once again that the extensions and truth conditions of our thought and talk are governed more by our false quintessentialist mindset than by any "scientifically discoverable" essentialism, contrary to what philosophers often suppose.[28]

2.4.1. Essentialism and Compounds

Let us begin with the case of water, which we are told, is "identical" to H_2O. One initial observation that has not been properly assimilated is that "H_2O" is *not* in fact a microstructural description (Needham, 2000, 2002, 2011; van Brakel, 1986). It is simply a compositional formula, specifying the proportion of hydrogen to oxygen that is to be found in the substance. It is not a specification of the molecular structure of water; in fact, it can apply to phases of water that *have no* molecular structure such as ice X, as noted below. Moreover, historically speaking compositional formulas were introduced and widely used at a time before atomism in chemistry was generally accepted; in this sense, compositional formulas do not even involve a commitment to the *existence* of atoms or molecules, the actual constituents of microstructure (Needham, 2000). Further, since compositional formulas only specify the proportion of elements in a compound, they do not even distinguish between *structural isomers*, which are chemically distinct compounds composed of the same proportions of elements. For example, the compositional formula "C_2H_6O" applies equally to ethyl alcohol and dimethyl ether, but these two compounds are distinct along any number of chemically important dimensions. As it happens, water has no structural isomers, however many, many compounds do. (And many compounds have many, many structural isomers. For example, in the case of large organic molecules—even ones that consist only of carbon and hydrogen—the number of distinct compounds corresponding to a given compositional formula can number in the millions and greater (Smith, 2011).) A somewhat better candidate for capturing the 'microstructure of water' would be a structural formula—a formula which would specify to some approximation how the relevant molecule is structured, for example by indicating that the two hydrogen atoms are each bonded to the central oxygen atom. There are numerous ways of representing structural formulas, many of them

[28] This next section draws on the excellent work of a number of philosophers of chemistry, including Michael Weisberg, Joseph LaPorte, Jaap van Brakel, and especially on the exemplary work of Paul Needham, whose papers should be mandatory reading for philosophers working on natural kinds. The presentation and many of the illustrations are my own, however.

graphic in nature, so as to illustrate the geometry of the bonding patterns. For simplicity, I will use 'H–O–H' as an abbreviation for a more informative structural description of the molecule.

The real complications arise when we learn that only in the gas phase does water primarily consist of such H–O–H molecules. That is, the gas phase of water can be reasonably modeled (though with some idealization) by supposing that there are a large number of discrete, separate H–O–H molecules, but this is not true for the liquid phase, nor for the many ice phases. For example, there are at least fifteen different forms of ice that have been observed experimentally, each formed under different combinations of temperature and pressure. (And more are theoretically projected, including a metallic phase of ice!) Interestingly, one form of ice, ice X, has no molecular structure at all; its microstructural arrangement is such that there is no distinction between the *intra*-molecular bonds and the *inter*-molecular bonds. Rather ice X is an atomic solid, composed of hydrogen and oxygen atoms arranged in a particular lattice structure. It does not contain any molecules to be described, so the question of *how* to describe its molecules does not arise. It still contains hydrogen and oxygen in a 2:1 ratio, however, and so falls under the compositional formula "H_2O," despite lacking any molecular structure.

In the liquid phase, there is molecular structure, but unlike the gas phase, only a sub-portion of the molecules are H–O–H molecules. For example, some of the H–O–H molecules dissociate into H + and OH– ions. Some of these ions attach to H–O–H molecules to form complex ions, for example H_3O + and $(H_2O)OH$– (since the complexity of the molecules is rapidly increasing, I will use condensed structural formulas such as these). Other molecules bond together to form polymers (i.e. chains of repeating structural units) of arbitrary length: two H–O–H molecules will combine to form a $(H_2O)_2$ molecule, which can combine with another H–O–H molecule to form $(H_2O)_3$, and so on and so forth. The polymers and the ions can further combine, so in a given sample of liquid water there may be molecules of the form $(H_2O)_nH$ + and $(H_2O)_nOH$– for any reasonable n. The patterns of disassociation and bonding in liquid water happen continuously—what is a polymer one moment may disassociate into smaller parts the next, and so on (see Needham, 2000, 2011 for more discussion).

This is the beginning (and just the beginning!) of a serious description of the microstructure of water. The idea that "H_2O" constitutes a description of water's microstructure may have its roots in the incorrect idea that the liquid and solid phases of water (and other substances) are basically like the (still somewhat idealized conception of the) gas phase—which we may say is composed by H–O–H molecules. That is, the idea may persist that the liquid and solid phases have the same microstructural composition, only with the molecules staying closer together in these phases. In fact the important manifest properties of water—for example its relatively high boiling point, which ensures it is liquid at standard temperatures and pressures and therefore drinkable; the fact that ice is less dense than liquid water, which means fish can survive the freezing of lakes by swimming lower; and so on—are due specifically to that fact that liquid and solid water do not simply consist of H–O–H molecules, but rather a variety of dynamically related complex molecules.

In light of this complexity, what are we to make of the very idea of a microstructural essence for water and other chemical kinds? As a general rule, a compositional formula like "H_2O" does not provide a good candidate for capturing the essence of a substance. For example, recall that the compositional formula "C_2H_6O" applies equally to ethyl alcohol (condensed structural formula "C_2H_5OH") and dimethyl ether (condensed structural formula "$(CH_3)_2O$)", but these two compounds are quite dissimilar (and certainly considered distinct by chemists), and we should not want the result that they are somehow to be said to share a microstrucural essence, and so be the same substance. In general, a compositional formula does not specify a unique chemical kind; in fact, in some cases the same compositional formula applies to *millions* of distinct chemical kinds (Smith, 2011). But on the other hand, relying on a structural formula to specify the composition of the constituent molecules will also be problematic as general strategy, as the case of water illustrates—if we attempt to insist on something to the effect that *water is essentially composed of H–O–H molecules*, then we will derive the result that water exists only in the gaseous phase. And certainly, no description of molecules will suffice to characterize ice X, since it does not contain molecules.

2.4.2. The 'Same Substance' Relation

Perhaps the solution will be to examine the microstructure of water and other chemical kinds in even greater detail, and hope that such examination will uncover the elusive essence of chemical kinds. Returning to the case of liquid water (which, remember, was supposed to be the easy case, the poster child for Kripke/Putnam essentialism!), we can begin by noting that the rates and patterns of polymer formation and ionic disassociation are sensitive to the temperature and pressure. Thus, if we take seriously the idea of deriving detailed microstructural descriptions, water at 4 degrees C will have a considerably different microstructure than water at 98 degrees C. Any microstructural specification of the essence of water would have to take account of this variation without delivering the result that these are different substances. As Needham puts it,

> In view of the sensitivity of the distribution of particular determinate forms of the many determinable features to prevailing conditions, the actual description of the relevant microscopic structure—which must be sufficiently relaxed to accommodate all these variations without including too much—would be a truly daunting task . . . But first and foremost, it should be asked why this vast range of microscopic structures is associated with one and the same substance kind, rather than a genus of related substances. (2000, p. 20)

Given the extensive variation at the microscopic level, is water even a single substance? And more generally: what criteria, if any, does chemistry give us for determining *how many* different substances are present in a given sample—or put differently, does chemistry give us a way of determinately answering Putnam's question 'is x the *same substance* as y'? We have seen that neither sameness of compositional formula nor sameness of molecular structure provides a satisfying answer to the question (because of the possibility of isomers and the possibility of massive variation on the microstructural level respectively). As I understand Needham's work, the real answer is two-fold: 1) such criteria for sameness are far more readily found at the macroscopic level than at the microscopic level, and 2) there is nonetheless no unique and privileged way of answering the question at either level.

Put in a slightly simplified way, the point is this: When it comes to manifest kinds, chemists typically find unexpected complexity and variability in the constituents of the examples of the kinds in question, and they are fully prepared to note this complexity and variation without any surprise, because they treat the manifest and macroscopic features as partly determinative of what it is to be, for example, water or gold—contrary to what Kripke and Putnam envisaged.

Chemistry's partial reliance on and respect for the macroscopic and the manifest is at odds with the Kripke/Putnam project.[29] Consider their repeated reference to common microstructure as wholly determinative of the extension of our natural kind terms. Consider, for example, Kripke in the following passage:

> what I would have wanted to do would have been to discuss the part about gold being a metal. This, however, is complicated because first, I don't know too much about chemistry. Investigating this a few days ago in just a couple of references, I found in a more phenomenological account of metals the statement that it's very difficult to say what a metal is. (It talks about malleability, ductility, and the like, but none of these exactly work.) On the other hand, something about the periodic table gave a description of elements as metals in terms of their valency properties. This may make some people think right away that there are really two concepts of metal operating here, a phenomenological one and a scientific one which then replaces it. This I reject. (1980, pp. 117–8)

The thought is that the valency properties or "something like that" is what makes things deserve the name of metals, rather than the macroscopic features of malleability, ductility and the like. In fact, there are at least two scientific

[29] Putnam, especially in his later work (but also in "The Meaning of 'Meaning"' p. 239), does not seem to take such a consistently hard line on this, and allows instead that our interests also matter when determining whether one thing counts as the *same substance* as another. See Hacking (2007a) for a helpful discussion. Again, here and throughout the paper, I use "the Kripke/Putnam project" to denote the one that is rooted in *Naming and Necessity* and "The Meaning of 'Meaning" ' as it is widely interpreted—that is, the view that has been so historically influential in philosophy.

concepts of metal in play in chemistry—one denoting a set of elements in the periodic table, the other applying when a solid of any atomic constitution displays certain macroscopic properties—as is illustrated by the hypothesis that there exists a metallic phase of ice. This latter is characterized in macroscopic terms, but is nonetheless an important scientific notion, despite Kripke's suggestion to the contrary. This dichotomy between the "scientific" and the "phenomenological"—especially if the latter includes notions such as malleability and ductility—is artificial and cannot be sustained. Indeed, there is an interesting further question as to whether the elements that are called metals are so-called because they *typically* form metallic solids at normal temperatures and pressures, where what counts as a metallic solid is "phenomenologically" determined, that is determined by such properties as malleability, ductility, and the like.

Returning to the question of the *same substance* relation, one promising place to look is what is known as Gibbs' phase rule, which relates 1) the number of independent substances in a closed system (thus offering a way to 'count substances'), 2) the number of phases in the system (e.g. solid, liquid, etc.), and 3) the number of 'intensive variables' (e.g. temperature, pressure), which can be varied independently. The phase rule does deliver the result that water is a single substance.[30] However, it must be understood that Gibbs' rule counts *independent* substances. If multiple substances are related to each other within

[30] The phase rule is as follows:

$c - f + 2 = V$

If $V \geq 0$, then the system will be able to be in equilibrium under at least some conditions. Suppose $c = 1$—that is, that we have but a single substance in the system—then the phase rule allows us to make empirical predictions about how that substance will behave with respect to its different phases, as a function of temperature and pressure. For example, if we have but a single substance in the closed system, Gibbs' phase rule lets us derive that there will be a unique combination of temperature and pressure at which three phases of the substance will be at equilibrium—that is, there will be a specific temperature and a specific pressure at which, e.g. solid, liquid, and gas phases can all co-exist in equilibrium ($V = 0$, so no intensive variables can be varied; this is sometimes known as the triple point). The phase rule also lets us predict that, if $c = 1$, there will be a range of temperature and pressure pairs at which two phases (e.g. solid and liquid) can co-exist in equilibrium, but also that temperature and pressure cannot be varied independently. However, if again $c = 1$, but only one phase is present, the temperature and pressure can each vary independently. To return to the question of whether water is a single substance or not, the answer is that, according to Gibbs' phase rule, it is. The behavior of water fits this thermodynamic profile of a single substance. As Needham illustrates:

> [S]uppose the two-phase quantity of water [liquid and gas] is confined at fixed temperature to a closed container fitted with a piston. Any attempt to decrease the pressure by raising the piston and increasing the volume available to the water will fail (as long as two phases remain), because the liquid phase will decrease in volume as matter goes into the gas phase to maintain the pressure. Similarly, attempting to increase the pressure by decreasing the volume will be thwarted by the volume of the gas decreasing as matter goes into the liquid phase to maintain the pressure. Continuing the processes of increasing or decreasing the volume will eventually lead to a single phase being formed (gas in the first case, liquid in the second), which is bivariant, so that pressure and temperature can vary independently. This behaviour is in accordance with the phase rule for a system for which the number of independent substances, *c*, is one (ms, p. 6).

This macroscopic criterion for individuating substances thus delivers the result that water is a single substance.

the system by a reversible chemical reaction, the phase rule will count fewer substances than seems correct. Consider, for example, what happens to a system that contains H_2O heated to around 2000K (Here I follow Needham, ms.). The H_2O begins to dissociate into H_2 and O_2 molecules, which in turn recombine into H_2O molecules. The phase rule, along with empirical observations about the behavior of the system, delivers the result that there is only one substance in the system, despite the fact that in general H_2, O_2, and H_2O are not the same chemical kind. What constitutes a distinct *independent* substance in a system is not in general the same as what constitutes a distinct substance simpliciter. The fact that the phase rule deems water to be a single substance could thus be interpreted in the following way: water is a conglomeration of distinct substances, but because they dissociate and recombine as they do, the distinct substances are not independent.

An alternative way of distinguishing substances is to consider the 'entropy of mixing'. Suppose we have two separate containers of gas sitting side by side, and we then remove the divider between them. If the two gases truly consist of the same substance, there will be no increase in entropy as a result of allowing them to combine. If, however, they are distinct substances then—even if they are highly similar—the entropy of the system (a measurable quantity) will increase (Needham, 2000). This turns out to be an exceedingly fine-grained test. For example, in addition to *structural* isomers (e.g. ethyl alcohol and dimethyl ether), there are also *spin* isomers, for example ortho- and para-hydrogen. A molecule of ortho-hydrogen is one in which the nuclear spins of the two H atoms are aligned parallel; a molecule of para-hydrogen is one in which the nuclear spins are anti-parallel. What is termed hydrogen gas and denoted by the formula "H_2" normally consists of a mixture of these two spin isomers (at a ratio of about 1:3 of para-hydrogen to ortho-hydrogen). The two spin isomers of hydrogen can be separated in a laboratory situation, however. Further, if a container of ortho-hydrogen is allowed to mix with a container of para-hydrogen, then the entropy of the system will increase; *by this test, there are thus (at least) two distinct substances in a normal sample of hydrogen.*

There are also spin isomers of water, ortho- and para-water, which are distinguished in the same way, depending on whether the spins of the hydrogen nuclei are aligned or not in the H–O–H molecule (and in H–O–H links in the polymer chains, and so on). Just as with hydrogen, normal samples of water contain both spin isomers, in a ratio of 1:3 (para to ortho). For various reasons, ortho- and para-water have proved more difficult to separate than ortho- and para-hydrogen, however several groups of researchers have recently succeeded in doing so (Kravchuk, Reznikov, Tichonov, Avidor, Meir, Bekkerman, and Alexandrovich, 2011; Tikhonov & Volkov, 2002). By the entropy of mixing test, ortho- and para-water count as distinct substances, suggesting that normal samples of water are in fact mixtures of two distinct substances. This disagrees with the results of the phase rule test, however, which suggests that water comprises a single substance.

So, are ortho- and para-water *really* distinct substances? The answer, in some sense, is clearly "yes." For example, efforts to separate them are clearly not fool's errands—attempts to separate a single thing into identical components. On the other hand, it is a distinction that is not generally made—water is normally treated as a single substance in chemistry. Tikhonov and Volkov, the first researchers to separate these spin isomers, found that the separated ortho- and para-water took 55 minutes and 26 minutes respectively to settle back into the normal 3:1 ratio when kept in a liquid state (far longer than they had expected), and estimated that ortho- and para-ice frozen at -18 degrees C would take several months to return to the normal ratio. They also note that "[t]he OP [ortho/para] separation procedure realized is quite straightforward and may occur in nature—in soil, atmosphere, living organisms, and cosmic objects. The scope and the role of this phenomenon are yet to be studied" (p. 2363).

It would thus not be totally beyond the realm of scientific possibility (though it may stretch the bounds) to imagine a version of the Twin-Earth thought experiment that compared ortho- with para-water. Would the visitors from Ortho-Earth be making an error when they first applied the word "water" to the substance on Para-Earth? Do these spin isomers bear the "same substance" relation to each other, or not? By some tests, yes (certainly both are H_2O); by other tests, no. There are scientifically useful ways of dividing up substances according to which they are the same substance, but there are other scientifically useful ways of dividing up substances according to which they are distinct. Nor are we awaiting the results of further investigation to settle which way of dividing up substances is *the scientifically real one*.

A similar example which is often discussed in the literature concerns isotopic variation (LaPorte, 2004; Needham, 2008; Weisberg, 2005). Two atoms are isotopes if they share the same atomic number (number of protons in the nucleus) but differ in mass due to the number of neutrons in their nuclei. Elements are standardly defined in chemistry by their atomic numbers, thus two distinct isotopes are nonetheless examples of the same element. Hydrogen atoms, for example, usually have a single proton and no neutron in their nuclei (protium); however, there are two naturally occurring stable hydrogen isotopes, one with a single neutron (deuterium), the other with two neutrons (tritium). Thus, protium, deuterium, tritium all fall under the same element kind (though of course they can obviously be distinguished from each other). Isotopic variation does not tend to affect the qualitative nature of the chemical interactions—if the most common variant of an element enters into a given chemical reaction, so will its isotopes. Sometimes isotopic variation can affect the speed of a chemical reaction, but much of the time the effect is negligible. Thus the addition, as it were, of a neutron—in contrast to the addition, say, of a *proton*—has little effect on an atom's chemical behavior. In this way, identifying elements with their atomic numbers regardless of isotopic variation is a very reasonable and useful way to classify elements and the compounds containing them.

In the case of hydrogen, however, isotopic variation has non-trivial effects—due in part to the fact that hydrogen atoms are so small (e.g. deuterium has roughly twice the atomic weight of protium). Consider, then, "heavy water," which has deuterium in place of protium. Is heavy water the same as regular water? By the most standard ways of categorizing compounds—which is insensitive to isotopic variation—they would count as the same substance. (Though of course by other ways of categorizing, they would be distinct—for example, there is an entropy of mixing associated with combining isotopic variants. Nothing in the observation that elements group isotopic variants together implies this is somehow a distinction without a difference.)

Imagine, then, a miniature version of Twin-Earth, in which a lake that contains heavy water is newly discovered, and let us pose the question of whether the substance in it is the same substance as the water encountered thus far. It should be clear from the foregoing that chemistry does not deliver a univocal answer to the question. Let us consider instead, then, how the members of population at large would be inclined to answer the question. While this paper in general cautions against reliance on intuition in philosophy, this would seem one occasion where we might nonetheless confidently "intuit" how this would go: the substance in this newly discovered lake would most certainly not be judged to be the same substance that comes out of our taps, etc. (LaPorte, 2004; Needham, 2008; Weisberg, 2005). This is for the straightforward reason that heavy water, when drunk in suitable quantities, causes sterility, neurological problems, and death. If approximately 25 percent of one's daily intake of protium oxide is replaced by deuterium oxide, one rapidly becomes sterile. If higher proportions of deuterium oxide are consumed, this soon leads to "acute neurological symptoms, liver hyperplasia, anemia, other symptoms, and eventually death" (Kushner, Baker, and Tunstall, 1998, p. 81). These effects, broadly speaking, are due to the fact that deuterium in heavy water exchanges with protium in organic molecules in the body. The bond between carbon and deuterium is considerably stronger than the bond between carbon and protium, which slows and otherwise disrupts a number of biochemical processes (Kushner et al., 1998).

I submit that we would not consult with chemists before concluding this substance was most certainly not the same substance as our water. Suppose an overenthusiastic philosopher should try to argue as follows: "in general isotopic variation has very little impact on chemical properties, and so elements are identified by their atomic numbers, and compounds by the elements in them, and so *it has been discovered* that the substance in this new lake is in fact *the very same substance* as the one we are used to drinking!" An apt response to the philosopher here might be: "well then, drink up!" The moral of this particular tale is that there are occasions on which chemistry does not deliver a univocal answer to the "same substance" question, yet the significance of the manifest properties of the substance nonetheless do provide such an answer.

We have seen thus far that chemistry does not provide a privileged *same substance* relation. Further, in the case of compounds, we also saw that there is no generally effective way of specifying the essence of the compound (even abstracting away from isotopism and spin isomerism)—simply specifying the composition of the compound in terms of its constituent elements is too broad (since it does not distinguish structural isomers), yet specifying the structure of the component molecules cuts too finely (since, e.g. water in its liquid phase is composed of many, many different types of molecules, while in its ice X phase it is not composed of molecules at all).

2.4.3. Essentialism and the Elements

What, though, of the elements? Certainly, the elements are more hospitable to essentialism than the compounds: in the context of chemistry, elements are quite clearly identified by their atomic numbers. (Again, though, as the case of isotopes illustrates, this does not clearly constitute a privileged *same substance* relation, as opposed to a *same element* relation.) Of course, for the Kripke/Putnam thesis to be correct, it is not enough that science provide the essences of kinds, it also has to be the case that these essences determine the extension of the relevant terms/concepts. Thus, an interesting question to explore is the extent to which the natural language terms/concepts for elements have their extensions determined simply by the relevant atomic number.

Consider, for example, Kripke's assertion that "a material object is (pure) gold if and only if the only element contained therein is that with atomic number 79" (1980, p. 138). Since the word "gold" is used in scientific contexts as the English name for the element Au, there is a sense in which something like this biconditional is unassailable as a necessary truth—this is the sense in which something is copernicum if and only if it has atomic number 112. The term "copernicum" is not a term of ordinary non-technical language; *copernicum* is not a manifest kind—rather it specifically names a certain element: the one with atomic number 112. That "gold" can have such a usage also is not something I would dispute. But Kripke intends to make a far more substantive claim: in fact he refers not to elements, but to material objects made of gold, and thus intends his statement to apply to the manifest kind *gold*—the kind which has played such an important role in our economic history. Gold's historical and economic significance depends, of course, on its manifest properties—for example, its lack of chemical reactivity make it very stable and free from corrosion, and its luster and general appearance make it desirable as jewelry and a store of wealth, while its malleability/ductility make it relatively easy for artisans and currency minters to work with.

But the manifest properties of an elemental substance depend on more than simply its atomic constituents: they also depend on how those constituents are arranged. When the atoms of a single element bond in different structural patterns, this is known as *allotropy*. As it happens, Au does not exhibit allotropy at normal temperatures and pressures, but rather forms only what is known as a face-centered cubic lattice, which accounts for many of the substance's important manifest properties, such as its ductility and malleability. It is not, however, *necessary* that Au atoms bond this way; for example, at very high pressures, Au atoms form instead a hexagonal close-packed lattice. The resulting substance is far less malleable and ductile than the face-centered cubic lattice of normal temperatures and pressures. If such allotropes of Au existed at normal temperatures and pressures, would we be so quick to assert that being composed of Au atoms is necessary and sufficient for falling under the manifest kind *gold*, despite lacking these economically significant properties? Or consider the element tin (Sn), which has in addition to its familiar shiny, metallic, malleable form, a non-metallic allotrope (stable below 13.2 degrees C), which is dull, brittle, and powdery. What if Au displayed such allotropic variation? Would such a dull, brittle, and powdery substance fall under the manifest kind *gold*?

One version of this question is a linguistic question: whether the word "gold" would be applied to such a counterfactual substance. A further question is whether this stuff would be considered the *same substance* as regular gold—for example, whether it would be assimilated into the same economic role as regular gold. Whatever the linguistic outcome would be in such circumstances, the economic outcome seems quite easy to predict: this counterfactual non-metallic Au would not play the same economic role as regular gold. (Of course if there were a cheap and easy way to convert the one to the other, that would change the economic role of the former, but only in a way that is parasitic on the role of the latter.) What of the linguistic question? My view here is that the linguistic question is less probative than it is often taken to be, in large part because—like most linguists who study lexical semantics—I believe that polysemy is a widespread and important phenomenon. That is, a single phonological form, for example */gold/*, may be associated with a range of related (sometime very closely related) but nonetheless distinct meanings. (Polysemy differs from lexical ambiguity, since in the latter case, the meanings are typically not at all related.) Thus, in the counterfactual circumstances we are considering, the possibility that the phonological form */gold/* might come to apply to the non-metallic allotrope would not show that the two were taken to belong to the same manifest kind. In this respect, economic behavior is more telling—since it requires one to "to put your money where your mouth is," as it were.

A real-world example of this form concerns the allotropes of carbon. Carbon has many allotropes, including graphite and diamond. It is an undeniable point that we do not consider graphite and diamond to belong to the same manifest kind, and the discovery that both consist of the 6th element of the periodic table does not alter this one bit (Johnston, 1997; LaPorte, 2004). It was a surprising chemical discovery that diamond contained only the humble carbon atom, but this did not somehow undermine diamond's economic role.

In the case of the 6th element, distinct phonological forms (at least in English) are used for each of the allotropes, and a further one (*/carbon/*) for the element itself. Had history gone differently, so that this 6th element was first discovered in the context of analyzing diamond, it may have been that this 6th element would have been named "diamond" also, paralleling the situation with gold. The surprising discovery would then have been that lowly graphite contained "diamond" atoms. This is a counterfactual history in which "diamond" is more polysemous than it actually is, not one in which the manifest kind *diamond* was discovered to include just anything composed of the 6th element.

For another illustration of polysemy as it applies to terms which have a use in which they pick out elements, let us consider the phonological form /*sodium*/. As background information, this phonological form applies to the 11th element of the periodic table. Pure macroscopic quantities of this 11th element can take the form of a metal, which is so reactive that it almost never occurs naturally, and must be covered with paraffin when stored in a laboratory. The phonological form /*sodium*/ also applies to this dangerous manifest kind. More commonly, however, the 11th element is found as a component of an ionic compound. Table salt is such a compound: it is composed of ions of sodium and chlorine. Sodium ions are typically ingested in the form of table salt, and play a number of extremely important roles in human physiology. for example sodium ions—in conjunction with potassium ions—help maintain blood pressure. The phonological form /*sodium*/ is again used without qualification in everyday discourse to pick out these ions.

Consider, then, a doctor's injunction to increase one's sodium intake, as a means of combating low blood pressure. The doctor here is not encouraging her patient to consume a highly reactive and toxic metal, though of course there is a perfectly standard use of /*sodium*/ that applies exactly to such a substance. Rather, the doctor is suggesting that her patient consume more *sodium ions*, probably in the form of table salt. (Note that for the halogens the two senses are distinguished phonologically: e.g. chlor*ine* vs chlor*ide*, where the latter refers only to ions.) In fact there are at least three related but distinct uses of /*sodium*/ in play here: the use that concerns ions as in "increase your intake of sodium", the use that concerns a metallic manifest kind as in "sodium is an explosive metal," and the use that concerns an element as in "sodium is the element with atomic number 11." It would however be a mistake—and in this case, a deadly one—to suppose that philosophical musings on chemistry establish that one of the latter two uses employs the *real* sense of "sodium." (Suppose someone makes this mistake and interprets his doctor's advice on increasing his sodium intake as the injunction to ingest a lump of sodium metal. One would not fancy the chances of a malpractice lawsuit pursued on the grounds that "sodium" is univocal and therefore applies to the lump of metal.)

Similar remarks apply to "oxygen." "Oxygen" can be used to pick out the 8th atomic element, and in this sense both O_2 gas and O_3 gas are composed of oxygen. The manifest kind *oxygen*—the one we must continually inhale to stay alive—includes only the former gas. The latter gas, ozone, is in an excellent sense *simply not oxygen*. (Inhaling ozone in sufficient concentrations rapidly leads to respiratory distress and eventually death.) When a paramedic declares that a patient needs oxygen, it is clear what she means, and what she means does not include ozone.

2.4.4. Beyond Chemistry: Quintessence and Chemical Kinds

When two manifest kinds are distinguished on the basis of their observable, macroscopic properties, could chemistry (or any other science) ever actually discover that they are one and the same? Of course we might discover they are closely related in some respect, and we may introduce a term that applies to both. Since polysemy is a widespread feature of the lexicon, it is possible that this umbrella term will be phonologically the same as the term for one of the two kinds. But if the two kinds genuinely differ in their important manifest properties, what discovery could lead us to discount this difference? For one thing, even if we allow a distinction between manifest properties and "scientifically significant" ones, the former do not float free of the latter—if there is a difference at the manifest level, there will be a difference at the "scientifically significant" level, and who is to say *that* difference is somehow unimportant?

Consider, for example, the fact that chemists generally use the term "water" in a phase-neutral sense. This does not mean that chemistry somehow obliges us to discount differences in phase—on the contrary, chemists are at times very interested in phase differences. What then of the term "water"? It is clearly used in a phase-neutral sense in most scientific contexts (such as when a chemist writes a paper), but in other contexts in a phase-specific sense (such as when the same chemist requests a glass of water). The term "water" is polysemous in this way.

Moreover, some of the best reasons for supposing that water in any sense comprises a single substance depend on its manifest and macroscopic properties, such as its behavior with respect to Gibbs' phase rule. In contrast to the Kripke/Putnam view, which supposes that when it comes to the determination of reference, the manifest and macroscopic is held wholly hostage to the microstructural and microscopic, the picture that emerges in chemistry itself is one on which macroscopic and manifest properties figure prominently.

The majority of examples considered thus far have been cases in which (a naïve and misleading conception of) chemistry might have led us to consider distinct manifest kinds to be one and the same. What then of the converse? Are there examples in which chemistry suggests that there is more than one kind present where there is arguably but one manifest kind?

If we take seriously the Kripke/Putnam emphasis on microstructure, then ortho-water and para-water provide such an example. Moreover, Joseph LaPorte (2004) elegantly and persuasively argues that the case of jade also fits this profile, and that the case of jade in China constitutes a historical Twin-Earth scenario that breaks in a way that is the opposite of what the Kripke/Putnam theory would predict.

Recall that there are in fact two chemically distinct compounds with extremely similar manifest properties—nephrite ($Ca_2(MgFe)_5Si_8O_{22}(OH)_2$) and jadeite ($NaAl(SiO_3)_2$). For approximately 5,000 years, nephrite played an extremely important economic and cultural role in Chinese society, and was more highly prized than even gold. In 1784, the Qianlong Emperor annexed Northern Burma. The region contained a number of mines that were rich in jadeite, a substance that the Chinese had not previously encountered. Jadeite and nephrite are exceedingly similar in their manifest properties, though not strictly identical; highly skilled artisans could tell the difference, since jadeite is slightly harder and heavier than nephrite. However, both substances share their most significant and prized manifest properties, for example their ability to sustain intricate carving. The Qianlong Emperor loved jade, and welcomed the newfound resources as more of the same substance. Jadeite was seamlessly assimilated into the same economic, artistic, and cultural role as nephrite, and the term "yü" was readily applied to jadeite and nephrite alike (Hacking, 2007b, LaPorte, 2004).[31]

This example again illustrates the significance of manifest properties in determining what counts as the *same substance* as something else. One should not, however, be mislead into supposing that manifest properties are *all* that matters. On this point, I believe that Kripke and Putnam were completely correct; however, I would interpret their insight in terms of our being, at heart, Quintessentialists. We do not believe that *looking like a K, feeling like a K*, and so on suffices for being a K. To truly be a K requires that one have the quintessence of Ks. On this point, it is instructive to note that, while both jadeite and nephrite have been embraced as true jade, other substances have been deemed false jade, despite having very similar appearances. That is, the manifest kind *jade* is not simply characterized in terms of superficial descriptive properties—the possibility exists for there to be exceedingly convincing cases of fake jade. (The Internet is replete with advice to buyers on how to avoid being conned in this way by unscrupulous jewelers.)

The situation may, I think, be characterized in the following way: we believe that there is a quintessence of jade, which is shared by both jadeite and nephrite, but not by other stones, no matter their appearance. (This leaves open whether we may believe that there are still minor quintessential differences between jadeite and nephrite—perhaps in the way that animal kinds below the basic-level are believed to differ in minor ways that are unimportant for general purposes. Compare Bengal tigers and Siberian tigers.) As the case

[31] Might this also be a case of polysemy? To best assess this question, it would be important to examine whether there were any notable economic or cultural distinctions drawn between the two. If none were drawn, then it would be reasonable to suppose that this is not a case of polysemy (at least not for most speakers of the language; for the artisans the situation may be different). I cannot rule out this possibility, but I do not think that chemical differences necessitate it either.

of jade suggests, quintessence does not map well onto chemical properties—though of course we may be inclined to suppose it does, just as we do in the biological case. A large difference in chemical properties does not entail a difference in supposed quintessence. I would further propose that the converse holds as, for example, the case of water and heavy water would illustrate: regardless of how chemically similar two substances may be, if one sustains life while the other ends it, these will be believed to have decidedly different quintessences.

As a final illustration of quintessentialist thinking, consider the distinction between natural and synthetic diamonds. Current technological advances have made it possible to produce diamonds in the laboratory which are chemically and physically identical to naturally occurring ones. In some cases it is possible for a trained gemologist with enough time and equipment to tell them apart, but even this is not fail-safe. Further, the ways in which they are distinguished often involve detecting imperfections that are more likely to be present in the *natural* diamonds than in the synthetic ones. It is certainly not possible for a consumer or even for the average jeweler to be able to tell them apart. Nonetheless, consumers are willing to pay far more for natural diamonds and, further, the overwhelming majority of consumers choose the natural ones.

These two kinds of diamonds are clearly distinguished in people's thinking: there are websites which offer advice on how to "get your fiancée to accept a synthetic diamond," and conversely websites in which people (including, we may suppose, the target fiancées) declare that they will "never settle for a synthetic diamond." Further, the qualifier "synthetic" is often replaced with a more derogatory term, such as "imitation," "simulated," or even "fake," which strongly suggests that these diamonds are not the *same thing* as the natural ones, but rather merely mimic their surface and chemical properties while being deeply and underlyingly different. As one jewelry website succinctly puts it, "Some customers may feel content to buy simulated diamonds, knowing that they could never afford a near flawless diamond. Others though, insist upon buying only genuine diamonds, as these are a more natural treasure and very rare, especially when compared to laboratory-produced items."[32]

The quintessentialist intuition here is that man-made diamonds *just aren't the same things* as natural ones. They are no more than forgeries, simulacra of the real things, whose natures are shaped deep in the ground under high pressure over many years. No lab can confer this nature on anything, because the nature does not simply reside in the chemical structure. The man-made stones simply do not have the quintessence of diamonds. Reflecting on chemical composition and bonding structure does not silence this complaint because

[32] <http://www.abazias.com/diamondblog/diamond-education/the-difference-between-real-and-synthetic-diamonds>. Accessed 24 November 2012.

our quintessentialist dispositions project differences even when there is no difference in the underlying chemistry.

3. CONCLUSION

We are Quintessentialists, and we have a tendency to interpret science through the lens of our quintessentialist intuitions. Thus, at first glance, it may seem to us that science confirms our quintessentialist intuitions, when in fact the opposite is true. Closer inspection of the scientific findings shows that—against the very natural and highly intuitive expectations of Kripke and Putnam—there are not good prospects for the discovery of microstructural biological or chemical essences that determine the references of our natural kind terms. There are cases in which members of the same species have quite different DNA and conversely ones in which members of different species have quite similar DNA (recall for example the cutthroat and the rainbow trout). There are cases of sameness of manifest kind with different microstructures (as with jade) and cases of different manifest kinds with the same microstructure (as with natural and synthetic diamonds). Further, neither biology nor chemistry delivers a privileged *same species* or *same substance* relation. The highly intuitive character of the Kripke/Putnam picture derives from our quintessentialist mindset; it does not reflect scientific or metaphysical facts about the world. If we did not have something like quintessentialist beliefs then we would not have the relevant Kripke/Putnam intuitions. We do, however, have such beliefs, and we have them despite the fact that the world does not conform to them. Our intuitions here reflect only facts about us, not facts about the deep nature of reality.

More generally, this discussion of one of the most intuitive views in philosophy illustrates the urgent need to examine the sources of our "philosophical intuitions." In particular, it is important to understand whether a set of intuitions may be due to an early-developing and deep-rooted bias to see the world in a particular way. Such a finding does not entail that the intuitions are misleading, but it does suggest that they will be extremely compelling *even if they are misleading*. Thus, such a finding gives us reason to subject the intuitions and any conclusions derived from them to further scrutiny. If a conclusion is supported solely by such intuitions, this may be good reason to remain skeptical concerning the conclusion.A question remains: what if quintessentialism not only implicitly shapes our basic beliefs but also our semantic intentions; what if in using natural kind terms we often do in fact intend to refer to kinds united by a common quintessence? Science shows there are not in general good candidates to be the common quintessences, so what happens to reference under this assumption? It is either much less determinate than the work of Kripke and Putnam led us to suppose or, if this is different, it is determinate only relative to a quintessentialist fiction that we all tend to share. A full exploration of this question must be left for another occasion.

REFERENCES

Allendorf, F. W., and Leary, R. F. 1988. Conservation and distribution of genetic variation in a polytypic species: The cutthroat trout. *Conservation Biology, 2*, 170–84.

Astuti, R., Solomon, G. E. A., and Carey, S. (2004). Constraints on conceptual development. *Monographs of the Society for Research in Child Development, 69*.

Atran, S., Medin, D. L, Lynch, E., Vapnarsky, V., Ucan Ek′ and Sousa, P. (2001). Folkbiology doesn't come from folkpsychology: Evidence from Yukatec Maya in cross-cultural perspective. *Journal of Cognition and Culture, 1*, 4–42.

Barker, M. (2010). Specious Intrinsicalism. *Philosophy of Science, 77(1)* 73–91.

Basch, S. H. (1973). The intrapsychic integration of a new organ: A clinical study of kidney transplantation. *Psychoanalytic Quarterly, 42*, 364–84.

Baum, P., Schmid, R., Ittrich, C., Rust, W., Fundel-Clemens, K., et al. (2010). Phenocopy—A strategy to qualify chemical compounds during hit-to-lead and/or lead optimization. *PLoS ONE, 5*, e14272.

Belk, R. W. (1990). Me and thee versus mine and thine: How perceptions of the body influence organ donation and transplantation. In J. Shanteau and R. J. Harris (eds.), *Organ donation and transplantation: Psychological and behavioral factors.* American Psychological Association: Washington, DC, 139–49.

Cole, E. R., Jayaratne, T. E., Cecchi, L. A., Feldbaum, M., and Petty, E. M. (2007). Vive la difference? Genetic explanations for perceived gender differences in nurturance. *Sex Roles, 57*, 211–22.

Coley, J. D., Medin, D. L., and Atran, S. (1997). Does rank have its privilege? Inductive inferences within folkbiological taxonomies. *Cognition*, *64*, 73–112.

Coley, J. D., Medin, D. L., Proffitt, J. B., Lynch, E., and Atran, S. (1999). Inductive reasoning in folkbiological thought. In D. L. Medin and S. Atran (eds.), *Folkbiology*. Cambridge, MA: Bradford, 205–32.

Devitt, M. (2008). Resurrecting biological essentialism, *Philosophy of Science*, *75*, 344–82.

—— and Kim Sterelny (1987). *Language and Reality*, Oxford, Basil Blackwell.

Diesendruck, G. (2001). Essentialism in Brazilian children's extensions of animal names. *Developmental Psychology*, *37*, 49–60.

—— and Gelman, S. A. (1999). Domain differences in absolute judgments of category membership: Evidence for an essentialist account of categories. *Psychonomic Bulletin and Review, 6*, 338–46.

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

—— (1993). *The disorder of things: Metaphysical foundations of the disunity of science*. Cambridge, MA: Harvard University Press.

Ereshefsky, M. (1998). Species pluralism and anti-realism. *Philosophy of Science, 65*, 103–20.

Ereshefsky, M. (2010). What's wrong with the new biological essentialism? *Philosophy of Science, 77*, 674–85.

Frazier, B. N., and Gelman, S. A. (2009). Developmental changes in judgments of authentic objects. *Cognitive Development, 24*, 284–92.

Gelman, R. (1987). Cognitive Development: Principles guide learning and contribute to conceptual coherence. Paper presented at the *Meeting of the American Psychological Association*, Division I, New York.

Gelman, S. A. (2003). *The essential child: Origins of essentialism in everyday thought*. New York: Oxford University Press.

—— and Coley, J. D. (1990). The importance of knowing a dodo is a bird: Categories and inferences in 2-year-old children. *Developmental Psychology, 26*, 796–804.

Gelman, S.A., Collman, P., and Maccoby, E. E. (1986). Inferring properties from categories versus inferring categories from properties: The case of gender. *Child Development, 57*, 396–404.

Gelman, S. A. and Markman, E. M. (1986). Categories and induction in young children. *Cognition, 23*, 183–209.

—— —— (1987). Young children's inductions from natural kinds: The role of categories and appearances. *Child Development, 58*, 1532–41.

Gelman, S. A. and O'Reilly, A. W. (1988). Children's inductive inferences within superordinate categories: The role of language and category structure. *Child Development*, *59*, 876–87.

Gelman, S. A. and Rhodes, M. (in press). "Two-thousand years of stasis": How psychological essentialism impedes evolutionary understanding. In K. S. Rosengren, S. Brem, E. M. Evans, and G. Sinatra (eds.), *Evolution challenges: Integrating research and practice in teaching and learning about evolution*. Cambridge: Oxford University Press.

Gelman, S. A. and Wellman, H. M. (1991). Insides and essences: Early understandings of the nonobvious. *Cognition, 38*, 213–44.

Gendler, T. S. (2007). Philosophical thought experiments, intuitions, and cognitive equilibrium. *Midwest Studies in Philosophy, 31*, 68–89.

—— (2008a). Alief and Belief. *Journal of Philosophy,105,* 634–63.

—— (2008b). Alief in action (and reaction). *Mind and Language, 23*, 552–85.

—— (In press). On the epistemic costs of implicit racism. *Philosophical Studies.*

Ghiselin, M. (1987). Species Concepts, Individuality, and Objectivity. *Biology and Philosophy, 2*, 127–43.

Gibson, G. and Wagner, G. (2000). Canalization in evolutionary genetics: A stabilizing theory? *BioEssays, 22*, 372–80.

Gil-White, F. J. (2001) Are ethnic groups biological "species" to the human brain?: Essentialism in our cognition of some social categories. *Current Anthropology, 42*, 515–54.

Gottfried, G. M. and Gelman, S. A. (2005). Developing domain-specific causal-explanatory frameworks: The role of insides and immanence. *Cognitive Development*, *20*, 137–58.

Graves, J. L. (2001). *The emperor's new clothes: Biological theories of race at the millennium*. New Brunswick, NJ: Rutgers University Press.

Greenwald, A. G. and Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review, 102*, 8.

Greenwald, A. G., Poehlman, T. A., Uhlmann, E., and Banaji, M. R. (2009). Understanding and using the Implicit Association Test: III: Meta-analysis of predictive validity. *Journal of Personality and Social Psychology, 97*, 17–41.

Hacking, I. (2007a). Putnam's theory of natural kinds and their names is not the same as Kripke's. *Principia, 11*, 1–24.

—— (2007b). The contingencies of ambiguity. *Analysis, 67*, 269–77.

Harris, R. N., Semtlisch, R. D., Wilbur, H. M., and Fauth, J. E. (1990). Local variation in the genetic basis of paedomorphosis in the salamander Ambystoma talpoideum. *Evolution*, *44*, 1588–603.

Hayward, C. and Madill, A. (2003). The meanings of organ donation: Muslims of Pakistani origin and white English nationals living in North England. *Social Science and Medicine, 57*, 389–401.

Heyman, G. and Gelman, S. A. (2000). Beliefs about the origins of human psychological traits. *Developmental Psychology, 36*, 663–78.

Hirschfeld, L. A. (1996). *Race in the making*: *Cognition, culture, and the child's construction of human kinds.* Cambridge, MA: MIT Press.

Holter, B. D. (2009). The ontology of species: A radically pluralistic perspective. Ph.D. thesis, Washington State University.

Hood, B. M. and Bloom, P. (2008). Children prefer certain individuals over perfect duplicates. *Cognition*, *106*, 455–62.

Hood, B. M., Donnelly, K., and Byers, A. (2011). Moral contagion and the negativity bias in attitudes towards potential organ transplantation. *Journal of Culture and Cognition.*

Hull, D. L. (1965). The effect of essentialism on taxonomy: Two thousand years of stasis. *British Journal for the Philosophy of Science*, *15*, 314–26 and *16*, 1–18.

Inspector, I., Kutz, Y., and David, D. (2004). Another person's heart: Magical and rational thinking in the psychological adaptation to heart transplantation. *The Israeli Journal of Psychiatry and Related Sciences, 41*, 161–73.

Jayaratne, T. (2001). National sample of adults' beliefs about genetic bases to race and gender. Unpublished raw data.

Jayaratne, T. E., Gelman, S. A., Feldbaum, M., Sheldon, J. P., Petty, E. M., and Kardia, S. L. R. (2009). The perennial debate: Nature, nurture, or choice? Black and White Americans' explanations for individual differences. *Review of General Psychology*, 13, 24–33.

Johnson, C. N. (1990). If you had my brain, where would I be? Children's understanding of the brain and identity. *Child Development, 61*, 962–72.

—— and Jacobs, M. G. (2001, April). *Enchanted objects: How positive connections transform thinking about the very nature of things*. Poster presented at the Society for Research in Child Development, Minneapolis, MN.

Johnston, M. (1997). Manifest Kinds. *Journal of Philosophy*, *94*, 564–83.

Kalish, C. W. (1998). Natural and artificial kinds: Are children realists or relativists about categories? *Developmental Psychology*, *34*, 376–91.

Kamp, H. and Partee, B. (1995). Prototype theory and compositionality. *Cognition, 57*, 129–91.

Keil, F. C. (1989). *Concepts, kinds, and cognitive development*. Cambridge, MA: MIT Press.

Keller, J. (2005). In genes we trust: The biological component of psychological essentialism and its relationship to mechanisms of motivated social cognition. *Journal of Personality and Social Psychology, 88*, 686–702.

Kitcher, Philip (1984), "Species", *Philosophy of Science,* 51, 308–33.

—— (2007). Does 'race' have a future? *Philosophy and Public Affairs*, *35*, 293–317.

Kolon, T. F. (2008). Disorders of sexual development. *Current Urology Reports, 9*, 172–7.

Kravchuk, T., Reznikov, M., Tichonov, P., Avidor, N., Meir, Y., Bekkerman, A., and Alexandrovich, G. (2011). A magnetically focused beam of ortho-water. *Science, 331*, 319–21.

Kripke, S. (1980). *Naming and necessity.* Cambridge, MA: Harvard University Press.

Kushner, D. J., Baker, A., and Dunstall, T. G. (1999). Pharmacological uses and perspectives of heavy water and deuterated compounds. *Canadian Journal of Physiological Pharmacology*, *77*, 79–88.

Lanie, A. D., Jayaratne, T. E., Sheldon, J. P., Kardia, S. L. R., and Petty, E. M. (2004). Exploring the public understanding of basic genetic concepts. *Journal of Genetic Counseling, 13*, 305–20.

LaPorte, J. (1997). Essential Membership. *Philosophy of Science*, 64, 96–112.

—— (2004), *Natural kinds and conceptual change*. Cambridge: Cambridge University Press.

Leslie, S. J. (2011). Essence, Plenitude, and Paradox. *Philosophical Perspectives, 25*, 277–96.

—— (submitted). Essence and Accident: An Empty Distinction.

Lewontin, R. C., Rose, S., and Kamin, L. (1984). *Not in our genes: Biology, ideology, and human nature.* New York: Pantheon Books.

Lizotte, D. J. and Gelman, S. A. (1999). *Essentialism in children's categories*. Poster presented at the Cognitive Development Society, Chapel Hill, NC.

Luhmann, C. and Coley, J. D. (2000). *Domain specific relations between typicality and absolute category membership.* Poster presented at Northeastern University College of Arts & Sciences Experiential Education Expo, May 2000.

Maglo, K. N. (2011). The case again biological realism about race: From Darwin to the post-genomic era. *Perspectives on Science, 19*, 361–91.

Mahalingam, R. (2003). Essentialism, culture, and power: Representations of social class. *Journal of Social Issues, 59*, 733–49.

Malt, B. (1994). Water is not H2O. *Cognitive Psychology, 27*, 41–70.

Mayr, E. (1982). *The growth of biological thought*. Cambridge, MA: Harvard University Press.

—— (1988). *Toward a new philosophy of biology: Observations of an evolutionist*. Cambridge, MA: Belknap Press of Harvard University Press.

—— (1991). *One long argument: Charles Darwin and the genesis of modern evolutionary thought*. Cambridge, MA: Harvard University Press.

Meyer, M., Gelman, S. A., and Leslie, S. J. (Submitted). My heart made me do it: Children's essentialist beliefs about heart transplants.

Meyer, M., Leslie, S. J., Gelman, S. A, and Stilwell, S. M. (in press). Essentialist beliefs about bodily transplants in the United States and India, *Cognitive Science.*

Murphy, G. (2002). *The big book of concepts*. Cambridge, MA: MIT Press.

Needham, P. (2000). What is Water? *Analysis*, 60, 13–21.

—— (2002). The discovery that water is H20. *International Studies in the Philosophy of Science*, 16, 205–26.

—— (2008). Is water a mixture?—Bridging the distinction between physical and chemical properties. *Studies in History and Philosophy of Science*, 39, 66–77.

—— (2011). Microessentialism: What is the argument?, *Nous*, 45, 1–21.

—— (Manuscript). The phase rule and the notion of substance. Stockholm University.

Nemeroff, C. and Rozin, P. (1994). The contagion concept in adult thinking in the United States: Transmission of germs and interpersonal influence. *Ethos, 22*, 158–86.

Newman, G., Herrmann, P., Wynn, K., and Keil, F. C. (2008). Biases towards internal features in infants' reasoning about objects. *Cognition, 107*, 420–32.

Newman, G. and Keil, F.C. (2008). 'Where's the essence?': Developmental shifts in children's beliefs about the nature of essential features. *Child Development, 79*, 1344–56.

Okasha, Samir (2002). Darwinian metaphysics: Species and the question of essentialism. *Synthese 131*, 191–213.

Pennell, W., and Barton, B. A. (1996). *Principles of Salmonid culture*. Netherlands: Elsevier.

Phillips, W., Shankar, M., and Santos, L. R. (2010). Essentialism in the absence of language? Evidence from rhesus monkeys (Macaca mulatta). *Developmental Science*, *13*, F1–7.

Phillips, W. and Santos, L. R. (2007). Evidence for kind representations in the absence of language: Experiments with rhesus monkeys (Macaca mulatta). *Cognition*, *102*, 455–63.

Putnam, H. (1975). "The Meaning of 'Meaning'." In H. Putnam, *Mind, language, and reality.* Cambridge: Cambridge University Press, 215–71.

Rhodes, M. and Gelman, S. A. (2009a). A developmental examination of the conceptual structure of animal, artifact, and human social categories across two cultural contexts. *Cognitive Psychology, 59*, 244–74.

—— —— (2009). Five-year-olds' beliefs about the discreetness of category boundaries for animals and artifacts. *Psychonomic Bulletin and Review, 16*, 920–4.

Rosch, E. H. (1973). Natural categories. *Cognitive Psychology*, *4*, 328–50.

—— (1975). Cognitive representation of semantic categories. *Journal of Experimental Psychology*, *104*, 192–233.

—— (1978). Principles of Categorization. In E. Rosch and B. B. Lloyd (eds.), *Cognition and Categorization*. Hillsdale, NJ: Erlbaum, 27–48.

—— Mervis, C. B., Gray, W. D., Johnson, D. M., Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology, 8*, 382–439.

Rosenberg, A. (1994). *Instrumental biology or the disunity of science*. Chicago: University of Chicago Press.

Rozin, P. and Nemeroff, C. (2002). Sympathetic magical thinking: The contagion and similarity "heuristics". In Gilovich, T., Griffin, D., and Kahneman, D. (eds.) *Heuristics and biases*: *The psychology of intuitive judgment*. Cambridge: Cambridge University Press, 201–16.

Salmon, N. U. (1979). How Not to Derive Essentialism From the Theory of Reference. *Journal of Philosophy*, *76*, 703–25.

Samarapungavan, A. and Wiers, R. (1997). Children's thoughts on the origin of species: A study of explanatory coherence. *Cognitive Science*, *21*, 147–77.

Sanner, M. A. (2001a). Exchanging spare parts or becoming a new person? People's attitudes toward receiving and donating organs. *Social Science and Medicine*, *52*, 1491–9.

—— (2001b). People's feelings and beliefs about receiving transplants of different origins—questions of life and death, identity, and nature's border. *Clinical Transplantation*, *15*, 19–27.

Sawin, P. B. (1932). Hereditary variation of the chin-chilla rabbit: In coat and eye color. *Journal of Heredity*, *23*, 39–46.

Shtulman, A. (2006). Qualitative differences between naïve and scientific theories of evolution. *Cognitive Psychology*, *52*, 170–94.

—— and Schulz, L. (2008). The relation between essentialist beliefs and evolutionary reasoning. *Cognitive Science*, *32*, 1049–62.

Simons, D. J. and Keil, F. C. (1995). An abstract to concrete shift in the development of biological thought: The insides story. *Cognition*, *56*, 129–63.

Smith, M. B. (2011). *Organic chemistry: An acid-based approach.* Boca Raton: Taylor & Francis.

Sober, Elliott (1980), "Evolution, population thinking and essentialism", *Philosophy of Science*, *47*, 350–83.

Solomon, G. E. A. and Johnson, S. C. (2000). Conceptual change in the classroom: Teaching young children to understand biological inheritance. *British Journal of Developmental Psychology*, *18*, 81–96.

Sousa, P., Atran, S. and Medin, D. L. (2002). Essentialism and folkbiology: Evidence from Brazil. *Journal of Cognition and Culture*, *2*, 195–223.

Springer, K. and Keil, F. C. (1989). On the development of biologically specific beliefs: The case of inheritance. *Child Development*, *60*, 687–48.

Stanford, P. Kyle (1995), "For pluralism and against realism about species", *Philosophy of Science*, *62*, 70–91.

Sterelny, Kim, and Paul Griffiths (1999), *Sex and Death*. Chicago: University of Chicago Press.

Stich, S. (2009). Reply to Sosa. In M. Bishop and D. Murphy (eds.), *Stich and His Critics.* Oxford: Blackwell, 228–36.

Taylor, M. G. (1996). The development of children's beliefs about social and biological aspects of gender differences. *Child Development*, *67*, 1555–71.

—— Rhodes, M. and Gelman, S. A. (2009). Boys will be boys; cows will be cows: Children's essentialist reasoning about gender categories and animal species. *Child Development*, *79*, 1270–87.

Templeton, A. (1999). Human Races: A Genetic and Evolutionary Perspective. *American Anthropologist*, *100*, 632–50.

Tikhonov, V. I., & Volkov, A. A. (2002). Separation of water into its ortho and para isomers. *Science*, *296*, 2363.

van Brakel, Jaak (1986), The Chemistry of Substances and the Philosophy of Mass Terms. *Synthese*, *69*, 291–324.

Waddington, C. H. (1942). Canalization of development and the inheritance of acquired characteristics. *Nature*, *3811*, 563–5.

Walker, S. (1999). Culture, domain-specificity, and conceptual change: Natural kind and artifact concepts. *British Journal of Developmental Psychology*, *17*, 203–19.

Waxman, S. R., Lynch, E. B., Casey, K. L., and Baer, L. (1997). Setters and samoyeds: The emergence of subordinate level categories as a basis for inductive inference. *Developmental Psychology*, 33(6), 1074–90.

Waxman, S. R., Medin, D. L., and Ross, N. (2007). Folkbiological reasoning from a cross-cultural developmental perspective: Early essentialist notions are shaped by cultural beliefs. *Developmental Psychology*. 43(2), 294–308.

Weinberg, J., Nichols, S., and Stich, S. (2001). Normativity and epistemic intuitions. *Philosophical Topics*, *29*, 429–60.

Weisberg, M (2005). Water is Not H2O. In D. Baird et al. (eds.) *Philosophy of Chemistry: Synthesis of a New Discipline*. New York: Springer, 337–45.

West-Eberhard, M. J. (2005). *Developmental plasticity and evolution*. New York: Oxford University Press.

Wiggins, David (1980), *Sameness and substance*. Oxford: Blackwell.

Wright, S. (1934). The results of crosses between inbred strains of guinea pigs, differing in their number of digits. *Genetics*, *19*, 537–51.

Zuckerkandl, E. and Villet, R. (1988). Concentration-affinity equivalence in gene regulation: Convergence of genetic and environmental effects. *Proceedings of the National Academy of Science*, *85*, 4784–8.