

FROM HUME'S DICTUM VIA SUBMERGENCE TO COMPOSITION AS IDENTITY
OR MEREOLOGICAL NIHILISM

By

Einar Duenger Bohn

Abstract: I show that a particular version of Hume's Dictum together with the falsity of Composition as Identity entails an incoherency, so either that version of Hume's Dictum is false or Composition as Identity is true. I conditionally defend the particular version of Hume's Dictum in play, and hence conditionally conclude that Composition as Identity is true. I end by suggesting an alternative way out for a persistent foe of Composition as Identity, namely mereological nihilism.

Composition as Identity is, roughly, the thesis that a whole is identical with all its parts.

Disambiguating: either (i) a whole is identical with all its parts collectively, but not with each one of its proper parts (if any) distributively, or (ii) a whole is identical with all its parts collectively and with each one of them distributively, or (iii) a whole is not identical with all its proper parts (if any) collectively, but it is identical with each one of its parts distributively.¹ *Hume's Dictum* is, roughly, the thesis that any two distinct contingent things are such that each one of them can exist without the other. An intrinsic property F is *submergent* iff it (i) is perfectly natural, (ii) is instantiated by a proper part of some composite object x, but (iii) does not locally supervene upon the class of intrinsic properties of x.²

In section 1, I prove the possibility of submergence based on the assumptions of composition as identity being false and a particular version of Hume's Dictum being true. In section 2, I argue that despite the perhaps ultimate lack of a formal contradiction,

submergence is incredible, and that the argument in its favor is best considered a *reductio* against one of its premises. The upshot is that if one is a friend of Hume's Dictum, one should also be a friend of composition as identity on pain of submergence. Conversely, if one is a foe of composition as identity, one should also be a foe of Hume's Dictum on pain of submergence. Interestingly, this goes against all metaphysicians who accept Hume's Dictum, but reject composition as identity. In section 3, I give a defense of the particular version of Hume's Dictum that my argument employed on the assumption that one is already onboard with some version of Hume's Dictum or other, and hence provide a push towards composition as identity for *any* friend of Hume's Dictum. I end by pointing out a potential problem with this solution, and briefly consider an alternative solution to my *reductio*, namely to reject that there are any composite objects at all, in which case we have a new and interesting argument on our hands for the necessity of mereological nihilism. Overall conclusion: any friend of Hume's Dictum should either accept composition as identity or mereological nihilism.

1. THE CASE FOR SUBMERGENCE

There are four *background* assumptions. First, I, at least for now, assume the possibility of there being at least some mereologically composite objects, i.e. objects with proper parts. Rejecting this assumption is what leads one by virtue of my argument to mereological nihilism. We'll come back to this point at the end of the paper.

Second, I assume there is a legitimate notion of *well-founded objective naturalness* (or something equivalent). The intuitive idea is that some concepts provide a better fit with objective reality than others. For example, the concept of being an electron

“carves reality at its joints” better than the concept of being grue. Furthermore, some concepts provide a *perfect* fit with reality, i.e. they tell us what reality is *really* like, most objectively speaking.³ To what extent this second background assumption is essential to the argument for submergence I am myself not sure of, but certainly the well-foundedness is not needed (a mere comparative notion is sufficient). It might be that the objectivity of the naturalness is also inessential. Perhaps a more subjective or pragmatic notion will do, but such territories are very foreign to me, so I will leave them aside for now.

The third background assumption is that there is a legitimate notion of *intrinsicity*. Intrinsicity is notoriously difficult to define,⁴ but the intuitive idea is that a property is intrinsic to its instantiator just in case its instantiation is not dependent on anything external to (“outside the skin of”) the instantiator. So, for example, while the mass and shape of something are presumably among its intrinsic properties, its weight and distance from the sun are presumably among its extrinsic properties.⁵ We also say that two things are *duplicates* iff they don’t differ with respect to any intrinsic property. It is essential to my argument that some such perfectly natural intrinsic properties can be instantiated by proper parts of something.⁶ As far as I can tell, some such notion of intrinsicity is needed for my argument to go through. In any case, if no such notion is legitimate, I am not the only one in trouble, so I will henceforth simply assume that we have such a legitimate notion without further ado.

Finally, the fourth background assumption is that there is a legitimate notion of *qualitative* properties. Again, this notion is notoriously difficult to pin down, but the intuitive idea is that a property (or relation) is qualitative just in case its most perspicuous

definition involves no particular individual.⁷ So, for example, while properties like *being red* and *having mass* presumably are qualitative, properties like *standing next to Karen* and *being identical to David* are non-qualitative. Important for present purposes is that *having a part* is also a qualitative property because though it involves some particular or other, it involves no one particular. That is, the indefinite article makes sure it is qualitative due to not picking out any one particular, only some particular or other. So, for example, while *having a part* is a qualitative property, but *having this part* is not. Likewise, *standing next to someone called Karen* is qualitative, but *standing next to Karen* is not. Any properties defined or derived from only qualitative properties by virtue of Boolean operators and quantifiers are also qualitative properties. So, for example, *having some part that is F* is a qualitative property, if *F* is. All my talk of properties is henceforth restricted to such qualitative properties. A bit of reflection after having gone through my argument will show that letting non-qualitative properties into my argument makes it easier to reject the possibility of submergence, so by making this last background assumption I make the case against submergence harder for myself.

Based on this background picture (or something close enough to it), the argument for submergence has only two further premises.⁸ The first premise being that mereological composition is not a form of identity. That is, a mereologically composite object is *ontologically distinct* – a different ontological constituent – from its proper parts, both individually and collectively. Let's state this first premise as follows:

1. A composite whole is not identical with its proper parts, neither individually nor collectively

This premise is almost universally accepted among contemporary metaphysicians.⁹ The premise is in fact a bit stronger than what is strictly speaking needed for my argument for the possibility of submergence. All that argument needs is that a composite whole is not identical with any one of its proper parts individually, leaving it open whether a composite whole is not identical with all its proper parts collectively. But I will later suggest that by accepting a version of composition as identity according to which a composite whole is identical with all its proper parts collectively, we have a plausible solution to the puzzle of submergence. So, I simply make the stronger assumption of premise 1 from the outset so as to more easily pick a premise for denial at the end. The stronger claim of premise 1 is in any case widely accepted among metaphysicians, and should as such be allowed from the outset for the dialectical sake of my overall argument.

The second premise is a version of *Hume's Dictum*. Consider two distinct chairs. It is metaphysically possible that one of the chairs exists without the other, or that one of them intrinsically changes without the other. That is, the two chairs are not necessarily connected such that they must (metaphysically) either both exist or both fail to exist, or for some intrinsic property, either both instantiate it or neither instantiate it.¹⁰ The intuition behind Hume's Dictum is that this possible existential and metaphysical disconnection of the two distinct chairs generalizes to other cases as well. The problem is of course to say more exactly to what extent it generalizes.¹¹

For present purposes, I need the following version of this principle:

(HD): for any two numerically distinct possible contingent concrete objects x and y such that $N_x x$ and $N_y y$, it is not the case that for all possible worlds w , if $N_x x$ in w , $N_y y$ in w .

where N_z is any fixed class of intrinsic properties (the “intrinsic nature”) of some possible contingent concrete object z . That is, we fix on a class of intrinsic properties of some possible contingent concrete object z , and we express the fact that these properties are instantiated by z by ‘ $N_z z$ ’. According to HD, no such intrinsic nature of x metaphysically necessitates an intrinsic nature of y , if x and y are two numerically distinct possible contingent concrete objects.

Note that along one axis, HD is a bit weaker than the way Hume’s Dictum is usually stated, according to which it is taken to also hold of things other than mere objects. So, along this first axis it should be fairly uncontroversial. But along another axis, HD is stronger than the way Hume’s Dictum is usually stated, according to which the notion of *mereological disjointness* takes the place of HD’s numerical distinctness. So, along this other axis it merits a further defense. I will give such a defense in section 3 below, but for now, let’s state the second premise of the argument for the possibility of submergence:

2. HD is true

From premise 1 and 2, our first desired conclusion logically follows:¹²

Proof: consider a composite object x with a proper part y , and assume that y instantiates the intrinsic, perfectly natural property F . By 1 and 2, it is possible that x exists with all its intrinsic properties intact without y doing so (by y either failing to exist or changing in its intrinsic nature), which means that y and its intrinsic nature do not locally supervene upon x and x 's intrinsic nature, which in turn means, by definition, that F is a submergent property. *Q.E.D.*

2. SUBMERGENCE IS INCREDIBLE

The possibility of submergence entails the possibility of there being a whole with the intrinsic property *having a proper part that is F*, for some perfectly natural intrinsic property F , but without any proper part of the whole being F . *Proof:* any duplication of the whole will have the intrinsic property *having a part that is F*, but by the possibility of submergence the intrinsic nature of the proper parts fail to (locally) supervene upon the intrinsic properties of the whole, so it is still possible that there is no proper part that is F . *Q.E.D.* So, as a toy example, it might be the case that my body has the intrinsic property of having an arm with mass m without any of my arms having mass m . But this result is incredible, so we most likely have a *reductio* on our hands.

One might think that there are some obvious cases of submergence. For example, consider a composite object a composed of two and only two proper parts b and c . Assume a has mass 4, and that b and c each has mass 2. Since mass is additive, the composite a might still have mass 4 even though we re-distribute the masses of b and c so that they have mass 1 and 3 respectively. This way we can keep an intrinsic property of the whole, but change the intrinsic properties of the parts. But this is not submergence as

I have construed it above. Submergence in this particular case would rather be this: assume the composite a has mass 4, and that b and c each has mass 2. Then duplicate a, so that it retains *all* its intrinsic properties, including the property *has a proper part that has mass 2*, but re-distribute the intrinsic properties of each of the parts, so that b and c has mass 1 and 3 respectively. Then the duplicated composite object a still has the intrinsic property *has a proper part that has mass 2* even though no one of its proper parts any longer has mass 2. Even worse, the duplicated composite object a neither has the intrinsic property *has a proper part that has mass 1* nor the intrinsic property *has a proper part that has mass 3* even though its proper parts has mass 1 and 3 respectively. But that's not just incredible, that's absurd!

Here is Jonathan Schaffer (2010:56) on submergence:

For submergence, the intrinsic properties of the proper parts, along with the fundamental relations between these parts, must fail to supervene on the intrinsic properties of the whole. This is impossible because (i) any intrinsic property of the proper parts ipso facto correlates to an intrinsic property of the whole, namely, the property of having-a-part-with-such-and-such-intrinsic-property, and (ii) any relations between the parts also correlates with an intrinsic property of the whole, namely, the property of having-parts-thus-and-so-related. Fix the whole, and all its parts are fixed.

This is of course no argument, but rather a mere statement of submergence being metaphysically impossible by virtue of the claim that the following biconditional is true:

(*Ipsa Facto*): necessarily, x has the intrinsic property *has a proper part that is intrinsically F* iff some proper part of x is *intrinsically F* (for some perfectly natural property F)

Now, *Ipsa Facto* seems indeed true, but it is hard to say exactly why. There is no formal contradiction in its denial, which entails: possibly, x has the intrinsic property *has a proper part that is intrinsically F* and no proper part of x is *intrinsically F*. The property *has a proper part that is intrinsically F* is a different property from the property *is intrinsically F*, so we simply don't get from the possibility of submergence alone any immediate formal contradiction of the form: x is F & x is not F . So why exactly is submergence impossible?

For present purposes, it would of course be best if we could derive a strict formal contradiction from the possibility of submergence.¹³ Now, I believe we can, but it involves additional commitments: we need the axiom schema for *lambda abstraction*:

$$3. [\lambda x(\Phi(x))]a \leftrightarrow \Phi(a)$$

The intuitive idea behind 3 is just this: consider any wff Φ , no matter how complex. Make a Fregean “gap” in it by removing one of its subject terms. What's left is a one-place predicate supposedly expressing a property. A thing satisfies that predicate just in case it has the one-place property expressed by it.

From 3, we get the desired result, namely that the possibility of submergence is inconsistent (letting F be a perfectly natural property):

Proof: assume for *reductio* that submergence is possible. Then it is possible that a has the property *has a proper part that is intrinsically F*, but no proper part of a is intrinsically F. That is, assuming F is intrinsic, possibly: (i) $[\lambda x(\exists y(y \ll x \ \& \ Fy))]a$ and (ii) $\sim \exists y(y \ll a \ \& \ Fy)$. By 3, we get the instance: (iii) $[\lambda x(\exists y(y \ll x \ \& \ Fy))]a \leftrightarrow \exists y(y \ll a \ \& \ Fy)$. By (i) and (iii), we get: $\exists y(y \ll a \ \& \ Fy)$, which directly contradicts (ii). *Q.E.D.*

We can also take the following route (letting again F be intrinsic):

Proof: as an instance of 3: $[\lambda x(\exists y(y \ll x \ \& \ Fy))]a \leftrightarrow \exists y(y \ll a \ \& \ Fy)$. By Necessitation: necessarily, $[\lambda x(\exists y(y \ll x \ \& \ Fy))]a \leftrightarrow \exists y(y \ll a \ \& \ Fy)$, which is *Ipsa Facto*, which in turn contradicts the possibility of submergence. *Q.E.D.*

The last proof shows that Schaffer's *Ipsa Facto* logically follows from 3, which explains his intuition that submergence is metaphysically impossible.

But some might accept that “ $\exists y(y \ll a \ \& \ Fy)$ ” can be true even though there really is no part of a that is F. That is, some might accept that there can be true existential claims “unwitnessed” in the ontology. For those, the above *reductio* will perhaps be considered unsound due to a rejection of a strong objectual reading of the quantifiers in play, or perhaps due to a rejection of a structural match between a true sentence and its truth-conditions.¹⁴ Others might reject the above *reductio* on grounds of rejecting the universal legitimacy of lambda abstraction, perhaps for nominalistic reasons, or because

they deny conjunctive properties. So, it would be good if we could come up with another explanation for why submergence is metaphysically impossible, or why *Ipsa Facto* is true, that does not rest on the above proofs.¹⁵

As far as I can see, the beginning of such an explanation lies in the denial of one of the two premises in the argument for submergence from section 1, i.e. in either denying Hume's Dictum or accepting composition as identity. The reasons for accepting Hume's Dictum also provide reasons for denying *Ipsa Facto* (since it involves a metaphysically necessary connection), especially so in combination with a denial of Composition as Identity. So, conversely, the reasons for denying Hume's Dictum also provide reasons that leave it open to accept *Ipsa Facto* (since you thereby allow metaphysically necessary connections). According to Lewis (1986:88), the principle behind Hume's Dictum is, roughly speaking, that "anything can coexist with anything else... Likewise, anything can fail to coexist with anything else." By this spirit, any whole and its parts should be able to coexist as well as fail to coexist, especially if the whole and its part are *two different things!* But then, by denying any one of our two premises, we might have the conceptual space needed for an explanation of why a whole and its proper parts might not always be such that they can coexist as well as possibly fail to coexist, and hence the space for an explanation of why *Ipsa Facto* is true after all, even without its denial entailing a formal contradiction.¹⁶

So, which one of the two premises is false: that Hume's Dictum is true or that composition as identity is false? Since the falsity of composition as identity is almost universally accepted, I suspect most will deny Hume's Dictum, or at least the version I have given of it, which involves numerical distinctness rather than mereological

disjointness. But this is a mistake. In the next section, I give a conditional defense of premise 2 of the argument for submergence, i.e. for HD, and hence a conditional defense for denying premise 1 of the argument, i.e. for denying that composition as identity is false.

But in any case, the first goal of this paper is now achieved, namely to show that there is a great tension (if not a formal contradiction) in the conjunction of Hume's Dictum and the negation of composition as identity. So, if one is a foe of composition as identity, then one should also be a foe of a more liberal version of Hume's Dictum on pain of the seemingly incoherent possibility of submergence. Or, conversely, if one is a friend of a liberal version of Hume's Dictum, then one should also be a friend of composition as identity on pain of the seemingly incoherent possibility of submergence. This is an interesting result in itself because, if submergence is incoherent, any foe of composition as identity should then either reject Hume's Dictum in all its glory or provide a principled version of it that involves a notion of 'distinct' weaker than numerical distinctness. But as I will now argue, the latter turns out to be a difficult matter.

3. FOR HUME'S DICTUM

There is no non-question-begging argument for Hume's Dictum. A good reason to accept it is perhaps that it provides us with the best handle on which possible worlds there are and which there aren't (Lewis, 1986). But if one is antecedently skeptical of Hume's Dictum, one is equally skeptical of this reason for believing it, so it will provide a tendentious argument, at best. So, Hume's Dictum is perhaps best accepted as a

fundamental metaphysical truth (as it is critically discussed in Wilson, 2010), or perhaps as a fundamental methodological principle (as it is argued for in McPherson, forthcoming). I opt for the former.

In any case, a full defense of Hume's Dictum is an enormous task that cannot be undertaken here and now. So, for present purposes, I simply assume that one does in fact, for some reason or other, accept it. Given that, I will argue that one should *at least* accept the particular version of it that I gave earlier, namely HD:¹⁷

(HD): for any two numerically distinct possible contingent concrete objects x and y such that $N_x x$ and $N_y y$, it is not the case that for any possible world w , if $N_x x$ in w , $N_y y$ in w .

where N_z is any class of intrinsic properties of some possible contingent concrete object z .

The restriction of HD to only hold of *contingent* objects is uncontroversial. If there are necessary objects, then of course they are necessarily connected to some other objects. For example, if numbers are necessary existents, I cannot exist without numbers because numbers have to exist, and hence there is a necessary connection between distinct things, namely a number and me.

The restriction of HD to only hold of distinct *objects* (and their intrinsic natures) might seem strange. Why not properties too (independent of their instantiators)? Well, we can allow HD to hold more broadly of any two distinct properties too, but my argument doesn't need to. All it needs is HD to hold of any two distinct possible contingent concrete objects, and the relationship between the intrinsic properties of *those two*

objects. If one accepts it in the case of any two properties as well, then some of the problems to be discussed will need other, presumably more controversial solutions than the ones I suggest. (I will indicate below what kind of solutions are available in those cases.) By admitting properties, we also face the problem of what to do with necessarily co-extensional, and hence necessary connections between instantiations of many properties such as being equilateral and being equiangular. But for purposes of the above argument for submergence, we only need HD to hold of two distinct objects (and their intrinsic natures), not any two distinct properties independent of whether they are instantiated by one or two objects. Whether it also holds of any two distinct properties can thus remain open for present purposes.¹⁸ Note also that it is not objectionably *ad hoc* to formulate the principle only in terms of distinct objects (and their intrinsic natures). The intuition behind it, i.e. the intuition we try to formulate by the help of HD, starts out with distinct objects and the way they are in themselves, independent of each other. The initial intuition does not start out with properties as such. We consider two ordinary objects and come to the conclusion that there is no good reason to think that either one of them cannot, metaphysically speaking, exist without the other, or intrinsically change independent of each other; after all they are two, not one, and intrinsic properties are independent of things other than the instantiator. The contrary claim that they must co-exist or both fail to exist, or cannot intrinsically change along some dimension independent of each other, is thus a mysterious and seemingly unjustified claim, and hence is best abandoned (at least until we are convinced otherwise). We don't initially start on this sort of a rationale with just two distinct properties, independent of whether it is one or two objects that instantiate them. For example, the properties *being red* and

being colored just cannot fully come apart in one and the same object: if it is red, it must be colored!

The restriction of HD to only hold of two distinct *concrete* objects might be argued to be *ad hoc*. But for present purposes the only place this matters, as far as I can tell, is in the special case of *impure sets*, which we will consider shortly. Note also that restricting HD to concrete objects doesn't mean Hume's Dictum as such isn't applicable to non-concrete objects; it just means that we need not take a stand on it for present purposes. My argument only needs HD to hold of concrete objects; whether the more general idea also holds of non-concrete objects can remain an open question for present purposes.

The most controversial part of HD is its interpretation in terms of numerical distinctness. The intuition behind HD is that any two distinct things can, metaphysically speaking, possibly co-exist and possibly fail to co-exist, or intrinsically change independent of each other. Usually, philosophers interpret this as not meaning numerically distinct, but rather something weaker, often meaning *mereologically disjoint*, not having any part in common. This way, the principle states, intuitively, that no two non-overlapping things are such that they must either co-exist or both fail to exist, and neither are they such that they cannot intrinsically change along some dimension independent of each other. And if one understands Hume's Dictum this way, then the above argument for submergence is blocked because the object in which the submergent property is instantiated is *not* disjoint from the object in which the subvenient properties are instantiated, namely the whole it is a proper part of. That is, with Hume's Dictum

understood in terms of non-overlapping things, it is silent with respect to necessary connections between a whole and its proper parts, which are overlapping things.

But Hume's Dictum thus understood is Hume's Dictum ill understood. Consider again the initial intuition behind Hume's Dictum. We start by considering two things and ask ourselves whether they *must* co-exist (whenever they exist) or both fail to exist, and whether each one of them can intrinsically change along some dimension independent of the other. We then find no good, or convincing rationale for *why* they must co-exist or both fail to exist, or cannot thus intrinsically change independent of each other. So we conclude that each of the two things *can* exist apart from the other, and can intrinsically change independent of each other. It is crucial to keep in mind that we are here considering a very liberal form of possibility, namely metaphysical possibility. We are not considering merely nomological possibility, i.e. only the possible worlds that share our laws of nature.

But with this liberal form of possibility in mind, any claimed metaphysical impossibility should entail some form of *incoherence*, if not a formal contradiction.¹⁹ And the incoherence should not be only in relation to other controversial metaphysical theses, but rather it should be a self-standing form of incoherence, or at least incoherence in relation to only more or less uncontroversial metaphysical theses. For example, we cannot deduce a formal contradiction from the claim that an object is red all over at time *t* and blue all over at time *t*, but it is nonetheless considered impossible due to it being, in some sense or other, incoherent in itself. Any claim of metaphysical impossibility should be able to show a form of incoherence close to this color case, if not a formal contradiction. This is not to say that the incoherence must be obvious, or that our search

for it is infallible, but only that until we have found some such incoherence, we have reason to think we are not dealing with a metaphysical impossibility. This is so because a mere claim of metaphysical impossibility is screaming out for an explanation: in lack of incoherence, *why* is it impossible? After all, we are working with a very liberal form of possibility. Due to its historical prevalence, we might want to call this liberal notion of possibility, the *traditional notion of possibility* (cf. the discussion of Hume below, and fn.23).

Now, until noted otherwise, let's assume that the thesis of composition as identity is false (i.e. premise 1, from section 1). That is, we assume that any composite whole is numerically distinct from each of its proper parts individually as well as from all its proper parts collectively. In slogan form: the whole is something other than its proper parts. Now (for "reductio") consider an arbitrary case of two numerically distinct, but overlapping objects *x* and *y* such that *x* and *y* are supposedly necessarily connected somehow. Why must *x* and *y* thus either co-exist or both fail to exist? It cannot be in virtue of them being in any way identical because *x* and *y* are by assumption not identical with each other and neither one of them is identical with its proper parts, and ipso facto at least one of them not with the part they overlap on. But if it isn't in virtue of identity, what is it that makes *x* and *y* necessarily connected? Why must they, in the very liberal space of metaphysical possibility, either co-exist or both fail to exist? The whole point of appealing to Hume's Dictum is to point out the mystery in such necessary connections. It just doesn't help to appeal to overlap, given that a whole and its proper parts are numerically distinct things. Overlap explains nothing, if composition/parthood is not

understood in terms of identity. The same holds if we consider the metaphysical possibility of each one intrinsically changing independent of the other.²⁰

Consider any principle of mereological composition: for any xx , xx composes some y if, and only if, xx satisfies Φ .²¹ Consider its negation: it is not the case that for any xx , xx composes some y if, and only if, xx satisfies condition Φ ; which entails: for some xx , xx satisfies Φ , but compose nothing or compose something, but doesn't satisfy Φ . Whence the incoherency? Why is the negation not metaphysically possible? It remains a mystery, and in particular a mystery anyone *already* onboard with any version of Hume's Dictum should not accept. After all, it goes against the whole point of appealing to Hume's Dictum to begin with.

It is worth noting that I am not arguing that HD is justified by an intuition alone. I don't believe that intuitions by themselves can do much justificatory work at all. Rather, I am arguing that we start out considering two distinct objects, ask about their metaphysical connections, and find no incoherence in the metaphysical possibility of one of them existing or intrinsically changing without the other, and hence conclude that they can metaphysically thus come apart from each other. The last move is justified by the fact that we are considering a very liberal form of possibility, namely metaphysical possibility, and as such should be able to find incoherence in impossibility. This liberal notion of possibility does not rest on a mere intuition of what is possible, but rather on what is the only *principled* or *non-ad hoc* way to draw the line between possibility and impossibility, namely incoherence. As such it principally generalizes beyond the naïve cases and intuitions we start out with.²²

In lack of incoherence, possibility is thus the default; impossibility must be shown. This, I claim, is built into the idea we are trying to formulate as Hume's Dictum. So, restricting the principle to only hold of non-overlapping things cast doubt on the initial motivation behind the principle, unless overlap is understood in terms of identity somehow. So, if we accept the principle at all, we should accept its liberal version in terms of numerical distinctness, i.e. HD.

It is worth comparing David Hume (1739-40:I,III,IV) himself on the matter, on the basis of which I have formulated HD above:

There is no object, which implies the existence of any other if we consider these objects in themselves, and never look beyond the ideas which we form of them. Such an inference wou'd amount to knowledge, and wou'd imply the absolute contradiction and impossibility of receiving any thing different. But as all distinct ideas are separable, 'tis evident there can be no impossibility of that kind.

The idea seems to be this: implication entails necessity; and the negation of necessity entails incoherence; but no such incoherence is entailed by the negation of the claim that an object is thus necessarily connected to some other object; so no object is thus necessarily connected to any other object. This is what we earlier called *the traditional notion of possibility*.²³ I have argued that the principled way to draw the line between possibility and impossibility is along the line of incoherence, and that the principled way to understand Hume's Dictum is then by understanding Hume's term 'any other object' as meaning any other *numerically distinct* object. Any weaker understanding of 'other' ('distinct') casts doubt on the initial motivation for the principle, and thus partly undermines the main reason for believing it. So, if you believe in Hume's Dictum, you

should believe it in terms of numerical distinctness (at least when it comes to contingent concrete objects), not mere mereological disjointness, or something weaker.

But there are of course objections to reading Hume's Dictum in terms of numerical distinctness. Stoljar (2007) and Wilson (2010) objects to such a reading by the fact that it contradicts many widely held positions in metaphysics. But by itself this is no objection. There is no reason to think that metaphysical principles like HD can stay neutral between all metaphysical positions. The question is rather whether the rejection is a benefit or a cost.

Both Stoljar (2007) and Wilson (2010) provide the case of a determinate and its determinable. For example, the property of being red is necessarily connected with the property of being colored in the sense that nothing can be red without also being colored. Nonetheless, an instance of being red is a numerically distinct thing from an instance of being colored, so this seems like a counterexample to Hume's Dictum understood in terms of numerical distinctness.²⁴

But this is of course no counterexample to HD in my argument, namely Hume's Dictum restricted to only hold of *objects* (and their intrinsic natures) and perhaps even to only hold of *concrete* objects. It is no counterexample to HD that an object *x* cannot be red without being colored because it is simply not a case involving two numerically distinct objects and their respective intrinsic natures, as required by HD.²⁵ Neither Stoljar nor Wilson provides an objection to this solution.²⁶

Wilson (2010) also provides the case of an object and its singleton set. For example, Socrates and the singleton of Socrates are presumably necessarily connected in

the sense that the singleton of Socrates cannot exist without Socrates. Nonetheless, Socrates and the singleton of Socrates are numerically distinct things, so this seems like a counterexample to HD, and as such HD is incompatible with what seems to be a metaphysical truth about impure sets and their members.

There are at least five solutions to this problem. One solution is to simply deny the existence of sets. To many Humeans, this might be precisely what they do on independent grounds anyway.

A second solution is to deny impure sets, but retain the pure sets of mathematics. All my argument needs is that HD holds of concrete objects, and hence remain silent on the pure sets of mathematics, which are abstract, not concrete.

Third, one might restrict HD further to only hold of *purely* concrete objects. A set of concrete objects is not itself a purely concrete object. Such an impure set is something in between a purely concrete object and a purely abstract object. While I can become sick from eating a rotten apple, presumably I cannot become sick from eating the singleton of that rotten apple, so there is something less concrete about impure sets than there is about their members. But if so, it is not a counterexample to HD that the singleton of Socrates cannot exist without Socrates because it is simply not a case involving two numerically distinct purely concrete objects. The case for submergence only needs a possible case of purely concrete objects.

A fourth solution is to deny that sets have qualitative intrinsic natures. A set – pure or impure – is fully individuated by its members, which in turn are (plausibly) mereologically disjoint from their set,²⁷ so an impure set and its members are not numerically distinct objects whose intrinsic natures are necessarily connected because

though they are numerically distinct, the set has no intrinsic nature. Hence, we have here no counterexample to HD.

A fifth and final solution, the most controversial one, but the one I personally favor of the six, is to deny that impure sets are really numerically distinct things from their members. By the axiom of extensionality, sets are individuated in terms of their members. As such it is plausible that being conceptualized as an impure set of concrete objects is just one out of many different ways of conceptualizing the plurality of objects that forms it. Another such way of conceptualizing the plurality is simply as being a plurality of objects (not a set). Yet another such way is as being the fusion of those objects, and yet another such way is as being the plurality of mereological parts that compose them. Being a certain set is thus a *relational* property of something, or some things, with a concept as the relational unit. (The corresponding notion of conceptualizing something in different ways through concepts can be as objective as one likes. That is, it need not make set theory subjective, or dependent on actual humans in any way.) As such we don't have a counterexample to HD because it is simply not a case involving two numerically distinct things with their respective intrinsic natures, but instead a case involving one and the same thing (or things) with respect to different subclasses of its (their) extrinsic nature(s).

Other putative counterexamples to HD come from cases of supervenience. For example, many hold that mental states supervene upon physical states; that evaluative/normative facts supervene upon descriptive/natural facts; that semantic facts supervene upon facts of usage; etc. In such cases we have necessary connections going

from the supervenience base to what supervenes upon it: no difference in the supervenient things without a difference in what they supervene upon.

My overall argument has been given on the assumption that we accept some form or other of Hume's Dictum, so I can appeal to this assumption at this point. And, *given Hume's Dictum*, any case of supervenience of a class of properties A on a class of properties B must be a case of a necessary relation between two classes of properties instantiated by *one and the same instantiator*. That is, given Hume's Dictum, we must either identify the things that instantiate A with the things that instantiate B, or reject the supervenience of A on B, all on pain of violating HD.²⁸ For a true friend of Hume's Dictum, this result should be no cost, but in fact a happy consequence. Supervenience of properties without identification of the things that instantiate them is a mysterious metaphysical necessity crying out for an explanation.²⁹

I anticipate a final objection to my argument.³⁰ One might accept Hume's Dictum with respect to any two numerically distinct objects, but reject that it generalizes to their intrinsic natures as well. That is, one might accept that any two numerically distinct objects can exist apart, but reject that they can always do so without changing their intrinsic natures. A whole and its proper parts might be a good such case at hand: they can exist apart, but some of their intrinsic properties must then change.

The objection results in an extremely weak version of Hume's Dictum, of minuscule metaphysical interest. Most proponents of Hume's Dictum accept something stronger, so it might just amount to rejecting it altogether. One also wants to draw a principled line between on the one hand, the cases where some x and y can existentially come apart, but not without intrinsic change, and on the other hand, the cases where they

can existentially come apart with no intrinsic change. But this seems impossible. The best suggestion that I can think of is that the principled line goes between on the one hand, what's "inside the skin" of something and on the other hand, what's "outside the skin" of something. In other words, the line goes between a form of overlap and disjointness. One might then claim that the former allows x and y to existentially come apart, but not without intrinsic change, while the latter allows it with no intrinsic change. But the "inside-/outside the skin"-talk is metaphorical talk, and "overlap/disjointness"-talk by itself explains no mysterious connections away. So, basically, the suggestion is just a disguised and unsupported claim that submergence is impossible, and hence no non-question-begging objection to my overall argument.

So, I conclude that there is (at worst!) a good case to be made for HD understood in terms of numerical distinctness as above. So, based on the argument from section 1, it follows that on the assumption that the thesis of composition as identity is false, there is a good case to be made for the possibility of submergence. But based on the reflections from section 2, submergence seems incoherent, if it isn't simply logically contradictory. So, by the logic of a *reductio*, we have a good case to be made for composition as identity. Of course, the thesis of composition as identity comes in various forms, each with its own pros and cons.³¹ But my argument is simply an argument for any thesis that rejects premise 1 from section 1 above by accepting that a composite whole is identical with all its proper parts, either individually or collectively. By accepting composition as identity, it seems one cannot apply HD to freely re-combine a whole and its parts, because a whole would be identical with its proper parts, either individually or collectively.³²

But there is a potential problem with this solution. One might object that, pace Baxter (1988b), composition as identity is the thesis that a whole is identical with *all* its parts collectively, not with *each one* of its parts individually, and hence HD can still be applied to a whole and its *proper* parts. If so, composition as identity doesn't save us from the incoherent possibility of submergence.³³

But I believe even a version of composition as identity according to which a whole is identical with all its parts collectively, but not individually, does save us from the incoherent possibility of submergence. The short story is this: according to composition as identity, or the particular version of it that I find the most plausible anyway, all mereological notions are *defined* and/or *analyzed* from a notion of identity,³⁴ so no mereological notion, including overlap and proper parthood, can be the appropriate form of numerical distinctness needed for HD in the argument to get submergence. But I am under no illusion here: this is a heavy explanatory burden still on the shoulders of composition as identity. Many will remain unconvinced.³⁵ This is not the place to defend composition as identity. I have in any case given a logically valid argument for the possibility of submergence, so one must deny a premise if one disagrees with the possibility of submergence. But if composition as identity fails, and we are onboard with Hume's Dictum, and HD in particular, what's left?

The most viable alternative to accepting composition as identity, I believe, is to accept both premise 1 (HD) and 2 (the falsity of composition as identity), but reject the first background assumption we made, namely that there are at least some composite objects.³⁶ This premise was needed to get the very notion of submergence up and running; after all, submergence is defined as a property of a proper part. So, given

Hume's Dictum, rejecting composition as identity, forces us to reject the possibility of composite objects. We then have a new and interesting *reductio* for mereological nihilism on our hands, the view that reality must consist of only simples, things without proper parts.³⁷ Personally, I am in fact tempted to go this way, if composition as identity fails to exempt proper parts and wholes from the strict rule of Hume's Dictum.³⁸

CONCLUSION

I first argued that if HD is true and composition as identity is false, then submergence is metaphysically possible, but submergence seems metaphysically impossible, so either HD is false or composition as identity is true. I then argued that if Hume's Dictum is true at all, then HD is true, which by the above argument entails that any friend of Hume's Dictum should accept composition as identity on pain of the seemingly incoherent possibility of submergence. Finally, I pointed out that another plausible escape route for the persistent foe of composition as identity, but friend of Hume's Dictum, is to accept mereological nihilism instead of composition as identity. In short, based on Hume's Dictum, I have presented an argument for either composition as identity or mereological nihilism. A defense of one of the latter theories must be left for another time, but the mere connections here shown between Hume's Dictum on the one hand, and composition as identity and mereological nihilism on the other, are, it seems to me, prominent avenues for future research.³⁹

NOTES

¹ Version (i) is probably what most people have in mind when thinking of composition as identity today. But version (ii) is found in Baxter (1988a, 1988b) and version (iii) is found in McDaniel (forthcoming).

² The term ‘submergence’ is taken from Schaffer (2010:56).

³ Personally, I favor an account of well-founded objective naturalness along the lines of Lewis (1983, 1986). For accounts of *fundamentality* that might do the needed job for present purposes, see e.g. Sider (2011), Williams (2010), Rosen (2010), Schaffer (2009), and Fine (2001) (though Rosen (2010) denies that fundamentality is well-founded).

⁴ For some heroic attempts, see Lewis (1986, 2001a), Vallentyne (1997), and Lewis & Langton (1998).

⁵ Weight, as opposed to mass, depends on the gravitational force present. Whether shape is intrinsic is somewhat controversial. See e.g. Skow (2007). Note also that in addition to intrinsic properties, there are presumably intrinsic *relations* and intrinsic *plural* properties as well.

⁶ For example, my left arm must be able to instantiate an intrinsic property such as *mass*.

⁷ On qualitative properties, see e.g. Adams (1979) or Lewis (1986).

⁸ Again, much of the background picture might be inessential to my argument, but it is best to be fully explicit about what I keep in the background. I leave it to the reader to figure out which parts of the background picture is not really needed.

⁹ The premise is usually taken for granted, but for explicit defenses of this premise, see van Inwagen (1994), Yi (1999), Merricks (2003:20-28), McDaniel (2008, 2010), Sider (2007, forthcoming,a), and Cameron (2012). For the minority rejections of the premise, see Baxter (1988a, 1988b), Lewis (1991:3.6), Armstrong (1997:17), Bohn (2011, forthcoming), Cotnoir (forthcoming) and Wallace (2011).

¹⁰ I said intrinsic *property*, not intrinsic relation. It might of course be that they thus stand in some intrinsic relation, e.g. *having the same color*.

¹¹ For a critical discussion, see Wilson (2010).

¹² The notion of local supervenience in play is standard: a class of instantiated properties *A* *locally supervenes* upon a class of instantiated properties *B* iff there can (metaphysically) be no difference in the instantiation of *A* without a difference in the instantiation of *B*. In short: any *B*-duplicate is an *A*-duplicate!

¹³ I take it that metaphysical possibility is a proper sub-space of logical possibility: there are logically possible worlds that aren’t metaphysically possible, but there aren’t metaphysically possible worlds that aren’t logically possible. For example, a ball being red all over and blue all over at the same time is logically possible, but metaphysically impossible. A formal contradiction is logically impossible, and hence metaphysically impossible. So, if submergence is logically impossible by entailing a formal contradiction, it is thereby metaphysically impossible, but if it isn’t logically impossible, it might still be metaphysically impossible. I take all this to be fairly uncontroversial.

¹⁴ For the latter case, see e.g. Williams (2010).

¹⁵ Though, personally, I rest content with the above proofs.

¹⁶ Arguably, Lewis (1991:3.6; 1993) rejected the premise that composition as identity is false. According to Bohn (2011), Lewis held a strong version of it, not the mere “analogical” version often attributed to him.

¹⁷ My argument is thus henceforth aimed at any friend of any version of Hume’s Dictum.

¹⁸ Personally, I am inclined to think it doesn’t hold for any two distinct properties due to some of the counterexamples below.

¹⁹ I take the notion of incoherence to be a primitive notion that is not mere formal contradiction. We need such a notion in mathematics (cf. Gödel’s Incompleteness Theorem), and elsewhere.

²⁰ But note that it is in this last case of intrinsic change the problem of submergence arises, if x and y stand in the proper parthood relation to each other.

²¹ If it is Universalism, let Φ be $xx=xx$; if it is Restrictionism, let Φ be your favorite restriction; and if it is nihilism, let Φ be *being a singleton plurality*.

²² In particular, it thus generalizes from the naive one-one cases of two disjoint objects, say two disjoint chairs, to less naive one-many cases, say that of a whole and all its parts (collectively and individually).

²³ The traditional notion of possibility, the idea that the negation of a metaphysical necessity entails incoherence, is indeed a traditional one. It is, for example, argued for by Leibniz, e.g. in his (1686). It can also be found in Descartes (1641). For a contemporary defender, see e.g. Rosen (2002, 2006).

²⁴ The same objection can be given in terms of the necessarily co-extensional properties *equilateral* and *equiangular*.

²⁵ The same reply is given in the case of equilateral vs. equiangular.

²⁶ Another solution, if one wants to extend HD to properties independent of whether they are instantiated by one or two objects, is to accept that a determinate and its determinables are not numerically distinct. For example, being red is not numerically distinct from being colored, but rather *partially* identical with being colored (cf. the discussion of composition as identity below). Then it is perhaps no counterexample to HD that an object x cannot be red without being colored because it is simply not a case involving two appropriately numerically distinct properties. Personally, I favor restricting Hume’s Dictum to concrete objects.

²⁷ Though see Caplan et al (2011).

²⁸ I imprecisely say we must identify the *instantiators*, or *things* that instantiate A and B, but see Bricker (2006), where it is much more rigorously formulated, and defended, in terms of a technical notion of *subject matters*. On the notion of subject matters, see also Lewis (1988a; 1988b; 2001b). Bricker (2006) points out that it is no accident that the friends of Hume’s Dictum are also the ones that refuse to accept supervenience without reduction.

²⁹ On this point, in addition to Bricker (2006), see McPherson (forthcoming).

³⁰ Thanks to Bruno Jacinto for pressing me to be explicit on this.

³¹ Compare Baxter (1988a, 1988b), Lewis (1991:3.6), Armstrong (1997), Bohn (2011, forthcoming), Wallace (2011), and Cotnoir (forthcoming).

³² This is no commitment to mereological essentialism. See Lewis (1991:3.5).

³³ This amounts to weakening premise 1 as discussed right after its statement in section 1.

³⁴ See Cotnoir (forthcoming) and Bohn (forthcoming).

³⁵ In fact, many will remain unconvinced of the very coherency of composition as identity. But Cotnoir (forthcoming), in particular, shows that such an attitude is no longer viable. Rather, the question is whether composition as identity can perform the duties it is set to do.

³⁶ I will not consider rejecting the other background assumptions since I am not clear to what extent my argument really needs them. But the next best alternative is perhaps to reject that we have an intelligible notion of intrinsicity that can do the work needed. But note that this will affect not just my argument, but any argument that needs intrinsicity.

³⁷ For a defense of nihilism, see Dorr (2005) and Sider (forthcoming,b).

³⁸ David Ingram pointed out to me that it might be possible to re-construct an argument for submergence analogous to mine in terms of plural quantification and plural predication without the background assumption that there are at least some composite wholes. I leave it as an exercise for the reader to try to work this idea out in sufficient detail, and consider what it means for my argument.

³⁹ Thanks to Ben Caplan, Aaron Cotnoir, Bruno Jacinto and Jonathan Schaffer. Thanks also to an anonymous referee for this journal for great comments. Thanks also to audiences at the University of St Andrews, University of Nottingham, and Lingnan University, where earlier versions of this paper were presented. Thanks also to Arché Research Center for hosting me while writing the last version of this paper.

REFERENCES

- Adams, R.M. (1979). 'Primitive Thisness and Primitive Identity,' *The Journal of Philosophy* Vol.76, No. 1, pp. 5-26.
- Armstrong, D. (1997). *A World of States of Affairs*. Cambridge University Press.
- Baxter, D. (1988a). 'Identity in the loose and popular sense,' *Mind* 97, pp. 576-582.
- Baxter, D. (1988b). 'Many-One identity,' *Philosophical Papers* 17, pp. 193-216.
- Baxter, D. & A. Cotnoir (eds.) (forthcoming). *Composition as Identity*. Oxford University Press.
- Bedau, M.A. & Humphreys, P. (eds.) (2008). *Emergence*. The MIT Press. Cambridge Massachusetts.
- Bohn, E. D. (2011). 'Commentary on "Parts of Classes",' *Humana.Mente* Issue 19.
- Bohn, E. D. (forthcoming). 'Unrestricted Composition as Identity,' in Baxter & Cotnoir (eds.) (forthcoming).
- Bricker, P. (2006). 'The Relation between General and Particular: Entailment vs. Supervenience,' *Oxford Studies in Metaphysics*, vol. 2, Oxford University Press.
- Cameron, R. (2012). 'Composition as Identity Doesn't Settle the Special Composition Question,' *Philosophy and Phenomenological Research* Vol. 84, Issue 3, pp. 531-554.
- Caplan, B., Tilman, C., & Reeder, P. (2010). 'Parts of Singleton,' *Journal of Philosophy* 107.10, pp. 501-533.
- Chalmers, D., Manley, D., & Wasserman, R. (eds.) (2009). *Metametaphysics*. Oxford University Press.
- Cotnoir, A. (forthcoming). 'Composition as Generalized Identity,' in *Oxford Studies in Metaphysics*.
- Descartes, R. (1641). *Meditations on First Philosophy*. As in Descartes (1984).

-
- Descartes R. (1984). *The Philosophical Writings of Descartes Volume II*. Edited by J. Cottingham, R. Stoothoff, & D. Murdoch. Cambridge University Press.
- Dorr, C. (2005). 'What we disagree about when we disagree about ontology,' in Kalderon (2005).
- Fine, K. (2001). 'The Question of Realism,' *Philosopher's Imprint* 1: 1-30.
- Gendler, T.S. & Hawthorne, J. (ed.) (2002). *Conceivability and Possibility*. Oxford University Press.
- Hume, D. (1739-40/2007). *A Treatise of Human Nature*. Edited by D. F. Norton & M. J. Norton. Oxford University Press.
- Kalderon, M. (ed.) (2005). *Fictionalism in Metaphysics*. Oxford University Press.
- Kallestrup, J. & Hohwy, J. (eds.) (2007). *Being Reduced: new essays on causation and explanation in the special sciences*. Oxford University Press.
- Kleinshmidt, S. (ed.) (forthcoming). *Mereology, Topology, and Location*. Oxford University Press.
- Leibniz, G.W. (1686). 'On Contingency,' in Leibniz (1989).
- Leibniz, G.W. (1989). *Philosophical Essays*. Translated by Roger Ariew and Daniel Garber. Hackett Publishing Company. Indianapolis & Cambridge.
- Leonard, H. S. & Goodman, N. (1940). 'The Calculus of Individuals and Its Uses,' *The Journal of Symbolic Logic*, Vol. 5, No. 2, pp. 45-55.
- Lewis, D. (1983). 'New work for a theory of universals,' in Lewis (1999).
- Lewis, D. (1986). *On the Plurality of Worlds*. Blackwell Publishers.
- Lewis, D. (1988a). 'Statements partly about observation,' in Lewis (1998).
- Lewis, D. (1988b). 'Relevant implication,' in Lewis (1998).
- Lewis, D. (1991). *Parts of Classes*. Basil Blackwell.
- Lewis, D. (1993). 'Many but Almost One,' in Lewis (1999).
- Lewis, D. (1998). *Papers in Philosophical Logic*. Cambridge University Press.
- Lewis, D. (1999). *Papers in Metaphysics and Epistemology*. Cambridge University Press.
- Lewis, D. (2001a). 'Redefining 'Intrinsic',' *Philosophy and Phenomenological Research* Vol. LXIII, No.2, pp. 381-398.
- (2001b). 'Truth-making and Difference-making,' *Noûs* 35, pp. 602-615.
- Lewis, D. & Langton, R. (1998). 'Defining 'Intrinsic',' in Lewis (1999).
- McBride, F. (ed.) (2006). *Identity and Modality*. Oxford University Press.
- McDaniel, K. (2008). 'Against Composition as Identity,' *Analysis* Vol. 68, No.2, pp. 128-133.
- McDaniel, K. (2010). 'Composition as Identity Does Not Entail Universalism,' *Erkenntnis* 73, pp. 97-100.
- McDaniel, K. (forthcoming). 'Parthood is Identity?' in Kleinshmidt (ed.) (forthcoming).
- McPherson, T. (forthcoming). 'Ethical Non-naturalism and the Metaphysics of Supervenience,' *Oxford Studies in Metaethics*.
- Merricks, T. (2003). *Objects and Persons*. Oxford University Press.
- Rosen, G. (2002). 'A Study in Modal Deviance,' in Gendler & Hawthorne (eds.) (2002).
- Rosen, G. (2006). 'The Limits of Contingency,' in MacBride (ed.) (2006).
- Schaffer, J. (2009). 'On What Grounds What,' in Chalmers et al (eds.) (2009).
- Schaffer, J. (2010). 'Monism: The Priority of the Whole,' *Philosophical Review* Vol.119, No.1, pp. 31-76.

-
- Sider, T. (2007). 'Parthood,' *Philosophical Review* 116: pp. 51-91.
- Sider, T. (2009). 'Ontological Realism,' in Chalmers et al (eds.) (2009).
- Sider, T. (forthcoming,a). 'Consequences of Collapse,' in Baxter & Cotnoir (eds.) (forthcoming).
- Sider, T. (forthcoming,b). 'Against parthood,' *Oxford Studies in Metaphysics*.
- Skow, B. (2007). 'Are Shapes Intrinsic?' *Philosophical Studies* 13, pp. 111-130.
- Stoljar, D. (2007). 'Distinctions in Distinction,' in Kallestrup & Hohwy (2007).
- Vallentyne, P. (1997). 'Intrinsic Properties Defined,' *Philosophical Studies* 88, pp.209-219.
- van Inwagen, P. (1994). 'Composition as Identity,' in van Inwagen (2001).
- van Inwagen, P. (2001). *Ontology, Identity, and Modality*. Cambridge University Press.
- Wallace, M. (2011). 'Composition as Identity Part I & II,' *Philosophy Compass* Vol.6, Issue 11, pp. 804-827.
- Williams, J.R.G. (2010). 'Fundamental and Derivative Truths,' *Mind* Vol.119, No.473, pp. 103-141.
- Wilson, J. (2010). 'What is Hume's Dictum, and Why Believe It?' *Philosophy and Phenomenological Research* Vol. LXXX, No.3, pp. 595-637.