

# Fission, Fusion and Intrinsic Facts<sup>1</sup>

Katherine Hawley, University of St Andrews

## Abstract

Closest-continuer or best-candidate accounts of persistence seem deeply unsatisfactory, but it's hard to say why. The standard criticism is that such accounts violate the 'only a and b' rule, but this criticism merely highlights a feature of the accounts without explaining why the feature is unacceptable. Another concern is that such accounts violate some principle about the supervenience of persistence facts upon local or intrinsic facts. But, again, we do not seem to have an independent justification for this supervenience claim. Instead, I argue that closest continuer accounts are committed to unexplained correlations between distinct existences, and that this is their fundamental flaw. We can have independent justification for rejecting such correlations, but what the justification is depends upon much broader issues in ontology. There is no one-size-fits all objection to closest-continuer accounts of persistence.

## I

Here is a well-known problem about personal identity. Our persistence through time seems to be correlated with continuity of various kinds: psychological, physical, or both. But these continuity relations may hold between a single object existing at one moment and two distinct objects existing at a later moment: 'fission' cases seem possible, involving divided brains, amoeba-like body division or pressing the brain-state-transfer button twice. Fission is a problem for continuity-based accounts of personal persistence, because persistence through time is meant to be a matter of identity, and identity cannot hold between a single object and two distinct objects. Analogous problems arise for accounts of the persistence of artefacts, organisms, mountains and so on, since these are often expressed in terms of mereological,

---

<sup>1</sup> I am grateful to the Arts and Humanities Research Board and to my colleagues at St Andrews, for enabling me to take the research leave during which this paper was written. Material from the paper was presented at the universities of Aberdeen, Bristol, Leeds (at the Centre for Metaphysics and Mind), Rochester, Waterloo, Western Ontario and York (UK), and I am grateful to everyone who contributed to discussion. In particular, I would like to thank Bryan Frances, André Gallois, Hud Hudson, Fraser MacBride, Andrew McGonigal, Joseph Melia, Daniel Nolan, Eric Olson, Josh Parsons, Simon Prosser, Ted Sider, Gabriel Uzquiano, and the anonymous author of a very helpful referee's report.

material, functional, spatio-temporal and/or causal continuity, relations which are not necessarily one-one. For many kinds of thing, it is initially plausible to think that, for objects of that kind, some continuity relation is both necessary and sufficient for identity through time. The problem is that the logic of identity seems to forbid this.<sup>2</sup>

There are several well-known responses to this problem. First, we could argue that identity is not necessarily one-one.<sup>3</sup> Second, we could argue that there are two objects co-located before the ‘fission’, or that there is a single scattered or multiply-located object after the ‘fission’.<sup>4</sup> Third, we could give up the search for continuity-based persistence conditions, or give up the search for persistence conditions entirely.<sup>5</sup> Fourth, we could ‘cook up’ continuity relations which are guaranteed to be one-one.

Each of the first three options has its advantages and its disadvantages, but I won’t rehearse them here. Instead, my goal in this paper is to explore the prospects for the fourth option. Taking this option usually involves a ‘no-branching’, ‘closest continuer’ or ‘best candidate’ account of persistence (I won’t distinguish amongst these in this paper).<sup>6</sup> According to such accounts, for objects of a given kind, a certain type of continuity is both necessary for identity across time, and also sufficient where continuity does not ‘branch’. Roughly, *a* is identical to *b*, so long as *a* and *b* stand in the relevant continuity relations and there is no object, *c*, contemporaneous with *b*, which is also continuous with *a*. More precisely, the relevant continuity relation is typically a matter of degree, where reaching a minimum threshold is necessary for identity. If both *b* and *c* both achieve threshold continuity with *a*, then whichever achieves more continuity is identical with *a*; if they tie, then neither is identical with *a*.<sup>7</sup>

Closest-continuer accounts claim that a relation between objects can suffice for their numerical identity, while other possible objects seem to stand in the very same relation but

---

<sup>2</sup> Key literature includes Williams 1960, Parfit 1971 and 1984, Nozick 1981, and Shoemaker 1984.

<sup>3</sup> For example, Gallois 1998. Alternatively, we could adopt stage theory, claiming that persistence is not a matter of identity, and so not governed by the standard logic of identity. See Sider 1996, 2001, Hawley 2001.

<sup>4</sup> For example, Lewis 1976.

<sup>5</sup> For example, Merricks 1998.

<sup>6</sup> Sidelle 2000 explores an alternative version of the fourth option for certain sorts of thing.

fail to be identical because of different external circumstances. It all turns on the existence or nonexistence of a rival candidate. Many have found this claim objectionable.<sup>8</sup>

Why is it objectionable? A first thought is that closest-continuer accounts are untenably *ad hoc*: they result from modifying an account of persistence in response to a logical difficulty. But this objection is weak: it's not enough to point out that closest continuer accounts have a certain intellectual ancestry, unless we can explain why they are implausible in their own right. My main goal in this paper is to pin down this implausibility. I do not approach this task as a defender of closest-continuer accounts, but as a critic looking for a principled, theoretically-grounded basis for criticism.

Others who have attempted this task point out that closest-continuer accounts violate the so-called *only a and b* rule (a.k.a. the *only x and y* rule). According to David Wiggins, for example, "What we need, if identity is what we want to elucidate, is a criterion which will stipulate that for a relation R to be constitutive of the identity of *a* and *b*, *a*'s having R to *b* must be such that objects distinct from *a* or *b* are irrelevant to whether *a* has R to *b*. Let us call this the *Only a and b* condition...This condition is violated by the 'best candidate' proposal in all variants." (Wiggins 1980, 96) More recently, Wiggins writes "But the identity of *a* with *a*, of *b* with *b*, and of *a* with *b*, once we are clear which things *a* and *b* are, ought to be a matter strictly between *a* and *b* themselves. Let us call this...the *Only a and b* rule..." (Wiggins 2001, 96)

Consider how a closest psychological continuer account of personal identity handles putative fission. Suppose *a*'s body is destroyed, but *a*'s mental states are transferred to body B. The resulting person, *b*, is psychologically continuous with *a*. So, according to the relevant account of personal identity, *a* and *b* are identical. But suppose *a*'s mental states had been transferred both to body B and to body C. The person in body B and the person in body C would each be psychologically continuous with *a*, neither more closely than the other. So, according to the relevant account, neither of these would be identical to *a*. It looks as if,

---

<sup>7</sup> Parfit 1984 advocates such an account of people; Nozick 1981 advocates it for people, ships, and other material objects.

<sup>8</sup> For examples, see Wiggins 1980, Noonan 1985a, Heller 1987, Zimmerman 1999 and most of those who advocate alternative accounts of persistence. Garrett 1987, 1988, and 1990 responds to these objections.

according to this account of personal identity, whether *a* and the person in body B are identical depends upon what happens to body C; as Wiggins points out, this violates the *only a and b* rule.<sup>9</sup>

## II

As it stands, the *only a and b* rule does little more than reiterate that closest continuer theories are unacceptable. What justifies the rule? One thought is that it follows from the necessity of identity. After all, if the identity of *a* and *b* depends on the contingent (non)existence of a rival, then the identity is contingent: *a* and *b* are identical in rival-free worlds, and otherwise distinct. We are familiar with the perils of contingent identity.

Several authors have shown that closest continuer (CC) accounts can evade commitment to contingent identity.<sup>10</sup> Grant that, where fission does not occur, *a* is *b*, and that where fission does occur, the person in body B is not identical to *a*. We can uphold the necessity of identity if we accept that, where fission occurs, the person in body B is not *b*. Advocates of CC accounts would be committed to contingent identity only if they took ‘the person in body B’ to be a rigid designator, which they are not required to do.

It’s tempting to think that this response doesn’t really make the objection from the necessity of identity go away. After all, how can the reference of ‘the person in body B’ depend on what happens to body C? But critics of CC accounts would be ill-advised to pursue the objection from the necessity of identity further, because it doesn’t capture what’s intuitively objectionable about such accounts: a fan of contingent identity could reject a CC account of personal identity seemingly on the same grounds as the rest of us.

Such fans are typically motivated by problems of material constitution, claiming that, while a statue and its constituent lump are in fact identical, they could have been distinct, had the statue been made of different clay, and the lump been differently shaped. Identity between statue and lump is supposedly contingent upon their having various categorical properties.

---

<sup>9</sup> This objection can be run in the same way both for endurantists and for perdurantists, those who reject and those who embrace temporal parts. Matters are somewhat different for stage theorists, to whom I return shortly.

Now, someone who thinks both that identity may be contingent and that persistence is a matter of identity could nevertheless reject a closest-continuer account of persistence on seemingly the same grounds as the rest of us: how could *a*'s continued existence in body B depend upon what happens to body C? There seems to be something especially objectionable about making identity through time contingent upon the (non)existence of a rival.

This is not because facts about identity through time cannot in general be contingent on other facts. Pretty much any factor upon which the persistence of ordinary things might depend is contingent, with or without a closest-continuer account. It is always a contingent matter whether either psychological or material continuity occurs, and if such continuity is the basis of *a*'s persistence, then it is contingent how long *a* survives. Even on a straight continuity account, however, this doesn't introduce any contingency into whether *a* is identical to its later 'selves'.

Moreover, the objectionableness of treating persistence as contingent upon the (non)existence of a rival does not depend upon treating persistence as a matter of identity. Stage theorists claim that an object persists not by being identical with some three- or four-dimensional object existing at a later time, but merely by having a later counterpart.<sup>11</sup> Issues apparently about personal identity through time are really issues about the 'same person' counterpart relation. Confronted with symmetric fission, stage theorists have two options. They can claim that the pre-fission object has two post-fission counterparts (for the counterpart relation is unconstrained by the logic of identity). Or they can adopt a closest-continuer account, arguing that the pre-fission object has no post-fission counterparts. It seems evident that stage theorists should take the first option, and not the closest-continuer option. After all, how could whether some future person is 'me' depend upon what other candidates exist at that time? Whatever is objectionable about closest-continuer accounts of persistence remains objectionable whether we account for persistence in terms of identity or merely in terms of a counterpart relation.

---

<sup>10</sup> Nozick 1981, Salmon 1982, Shoemaker 1984, Noonan 1985a and 1985b, and Heller 1987.

<sup>11</sup> Sider 1996 and 2001, Hawley 2001.

It seems that what's objectionable about CC accounts is not a commitment to contingent identity, nor is it the fact that they treat persistence as contingent, and nor is it even a matter of their treatment of identity. Instead, it is something very specific about taking persistence to be rival-dependent: principles about the necessity of identity haven't provided any grounding for the *only a and b* rule.

Moreover, considering the 'same' objects in different possible worlds is misleading, for CC accounts make strange pronouncements even within a single possible world. Compare someone who survives transfer to a new body with someone else who allegedly perishes because the brain-state-transfer button is pressed twice. How can it be that the first survives, while the second does not? How can what is locally the same type of process count as the history of a single person in one set of external circumstances and the history of two sequential people in other external circumstances?

This last expression of the objection points away from theses about identity and towards some kind of supervenience thesis: the basic objection to CC theories emerges from an idea about the localness or intrinsicness of persistence facts. To substantiate the objection, we must clarify this idea and find some basis for it.

### III

In this section and the next, I will develop a principle about the localness of persistence which is violated by CC accounts but not by (apparently) more acceptable accounts. As we will see, such a principle exists, but it is quite far from being obviously true. It seems to owe whatever plausibility it has either to the original anti-CC intuition or else to some more fundamental principle: in later sections I will hunt down this more fundamental principle.

The 'localness of persistence' idea is roughly that if two regions match with respect to what's going on within them, they also match with respect to facts about the persistence of objects within them. CC accounts violate this supervenience principle, because they entail that whether or not a region like that beginning with *a* and continuing with the B body contains a

single persisting object depends partly upon what's going on outside the region.<sup>12</sup> The rough formulation needs polishing: first, what are the relata of the supervenience principle? second, which properties are in the supervenience base? third, which properties are supposed to supervene upon those base properties?

First, what are the relata? I have opted for regions, because it is hard to specify a substantive supervenience principle whose relata are objects or pairs of objects. 'If  $x$  and  $y$  match in respect  $R$ , then  $x=x$  if and only if  $y=y$ ' is trivially true for any consistent  $R$ . 'If pairs of objects  $(u,v)$  and  $(x,y)$  match in respect  $R$ , then  $u=v$  if and only if  $x=y$ ' is also easy to satisfy, even if we restrict  $R$  to features arising from intrinsic properties and exclude 'identity facts' from  $R$ . If  $u$  is identical to  $v$ , but  $x$  is not identical to  $y$ , then the pairs  $(u, v)$  and  $(x, y)$  are likely to differ intrinsically:  $u$  and  $v$  are of course duplicates whereas  $x$  and  $y$  are likely to differ from one another. The only possibility excluded by this principle is a pair of identical indiscernibles which intrinsically matches a pair of nonidentical indiscernibles.

We could specify a notion of time-bound properties of objects, or deal with temporal parts of objects. The former adds unnecessary complications, while the latter prejudices the debate: many of those who reject CC accounts also reject temporal parts. We could take processes as the relata of the supervenience relation, but these are often individuated by the objects which participate in them. So I will take the relata of the supervenience principle to be spatiotemporal regions, without claiming that regions are ontologically prior to the objects which occupy them.

Second, which properties of regions are in the supervenience base? The idea is to formulate a supervenience principle which will rule out closest-continuer accounts without deciding between, for example, psychological and bodily accounts of personal persistence, and without settling debates about Humean Supervenience.<sup>13</sup> So the supervenience base should be relatively rich, including all and only those qualitative properties whose instantiation is intrinsic to the region in question, facts about causal relations within the region, and any

---

<sup>12</sup> CC accounts are compatible with global supervenience principles: any two worlds which are alike with respect to continuity facts of the relevant sort are alike with respect to persistence facts about objects of the relevant kind. See Sider 1999.

<sup>13</sup> See, for example, Armstrong 1980, Zimmerman 1998, Lewis 1999, Hawley 1999, Callender 2001.

nonHumean facts intrinsic to the region.<sup>14</sup> Only the *qualitative* properties? This is intended to exclude facts explicitly concerning persistence or identity. Only the *intrinsic* properties? While it seems implausible that persistence facts should depend upon the (non)existence of rival candidates, it does seem more plausible that persistence facts should depend partly upon the laws of nature, which may well be extrinsic to a given region.<sup>15</sup> To sidestep this, let's restrict the supervenience principle to worlds which have the same physical laws.

Third, which facts are supposed to supervene upon this base of intrinsic, qualitative properties, within this circumscribed domain of worlds? Well, facts about persistence. Thinking in terms of persistence is preferable to thinking directly in terms of identity, partly because it encompasses the stage theory of persistence, and partly because it avoids issues about individual essences. The idea is not that whether a region contains Bill or Ben supervenes upon qualitative intrinsic facts about that region, but that whether or not a region contains a single persisting object supervenes upon that base.

So a candidate principle is this: two regions cannot differ in whether they exactly contain a single object without differing intrinsically.<sup>16</sup> But this principle is unsatisfactory, for two reasons. First, many philosophers believe that any matter-filled region of space-time exactly contains an object, and thus that if two regions match intrinsically, they match in whether they exactly contain a single object.<sup>17</sup> Such philosophers could maintain a closest-continuer account of persons, for example, without violating this supervenience principle. Even when fission takes place, an object exactly occupies the region occupied first by *a* and then by the B-body person; it's just that no person exactly occupies that region.

To rule out a closest-continuer view of persons (or ships, or whatever), we need a sortal-relative supervenience principle: two regions cannot differ in whether they exactly contain a single F without differing intrinsically. A sortal-relative principle will also be preferred by

---

<sup>14</sup> Of course, for the most part it is not the region but the object(s) within the region which instantiate the relevant properties. But I won't linger over this complication, since I will eventually abandon this supervenience strategy.

<sup>15</sup> Wasserman forthcoming.

<sup>16</sup> Take the aforementioned qualifications about laws, qualitative properties and so on as read. This talk of 'exact containment' is intended to be neutral between endurance and perdurance views of persistence; a reformulation would apply to stage theory too.

theorists who prefer disjunctive claims about the persistence of objects of various kinds to general claims about material persistence.<sup>18</sup> But going sortal-relative severely limits the scope of the principle, which cannot then be wielded against closest-continuer accounts of artefacts, or even biological things, whose kind-membership is not purely intrinsic. Yet CC accounts of persistence for these things are not attractive.

Second, even the sortal-relative version of this principle is too strong, because whether a region exactly contains an object of a given kind depends partly upon what's happening just outside its boundaries. David Hume could have been vaporized in 1775, a year before he actually died. Consider the Hume-shaped region which stretches from 1711 to 1775. Whether a region like this exactly contains a person depends upon whether the person it contains ceases to exist at its later boundary.<sup>19</sup>

A better sortal-relative formulation is this: two regions cannot differ in whether they each at least partly contain an F without differing intrinsically, where a region at least partly contains an F iff it is a proper or improper part of a region exactly containing an F. But discontinuous regions falsify this principle: whether a discontinuous region is part of a single-F region, or whether it contains a sequence of short-lived Fs may depend upon what happens in the 'gaps' (in neither case will there be direct causal connections between the various subregions). That's to say, intrinsically similar discontinuous regions could differ in whether either is part of a region occupied by a single object. We could just restrict our attention to continuous regions, but this would prevent us from objecting to CC accounts of the persistence of intermittently-existing objects, and if there were such objects, such accounts of their persistence would be objectionable.

How can we distinguish between a discontinuous region selected from a persisting F, and a discontinuous region exactly containing a sequence of distinct Fs? One way would be to allow facts about causal ancestry into the supervenience base. In the case of the

---

<sup>17</sup> Lewis 1986, Heller 1990.

<sup>18</sup> For example, Wiggins 1980, Lowe 1989, Hirsch 1993.

<sup>19</sup> Although it is contentious to claim that there are *objects* which are prevented from being Fs merely by their being large parts of Fs, it is entirely uncontentious to claim that there are sequences of events which could have been the entire life-story of an F, were they not parts of longer such sequences.

discontinuous selection from a single F, each distinct subregion bears an intimate and direct causal relation to a preceding F-containing subregion; this will not be the case for the discontinuous sequence of distinct Fs. Moreover, there are two further reasons to enrich the supervenience base to include not only causal relations within a region but also causal influences upon the region. First, an object's causal origins may help determine what sort of object it is, and thus what persistence conditions it obeys. Second, facts about causal ancestry are required in order to criticize CC accounts of fission whilst exonerating CC accounts of fusion. I turn now to showing why we should differentiate between fission and fusion in this way.

#### IV

Fission cases make problems for straight continuity accounts of persistence because an earlier object and a later one stand in an intrinsic relation that would suffice for identity were a rival candidate not simultaneous with the later object. The asymmetry between earlier and later is not obviously essential to the problem: the difficulty arises from the simultaneous existence of two objects each of which is continuous with a single object existing at another time. We might expect similar problems to arise involving two objects at an earlier time each of which is a candidate for identity with a single object existing at a later time.

How could such a case arise? Problematic fission cases are those in which we compare non-branching continuity with branching continuity. A problematic fusion case would compare the non-branching evolution of an earlier object into a later one with a process which differs only in that the later object has an extra causal 'source' in a further object also continuous with it. There are two ways in which this might happen.

First, the extra cause might overdetermine the state of the later object without changing it. Suppose I can persist through the destruction of my present body and the transfer of my brain states to a new body. Compare this with a situation in which both my brain states and those of my doppelganger are simultaneously transferred into the same new body. The post-transfer state of the person in the new body is intrinsically the same in each situation, as is my own pre-transfer state. My causal impact on the post-transfer person is overdetermined by

that of the doppelganger in the second situation. But there is a good sense in which causal relations between me-pre-transfer and the post-transfer person are the same in each situation, just as causal relations between earlier and later objects are unaffected by fission.

Second, the extra cause might instead replace an external cause of the later object without affecting the causal role of the earlier object. Suppose I can persist through the destruction of my present body and the accurate transfer of some of my brain states into a new body, despite the fact that others of my brain states are destroyed by the untimely use of a mobile phone. Compare this with a situation in which the transfer is not directly disrupted, but the brain-states of Keanu Reeves are simultaneously transferred into the same new body: the result is intrinsically just like that of the imperfect single transfer. In the second situation, the causal role of the disruptive telephone call is played by Keanu, who becomes my rival for identity with the later person. Again, the addition of the rival doesn't seem to affect intrinsic causal relations between me-pre-transfer and the post-transfer person, any more than causal relations between earlier and later objects are affected by fission.

How reliable are our intuitions about these 'fusion' situations? One sort of difficulty peculiar to intuitions about people is connected with demonstrative thought and memory. But there is also a difficulty concerning basic plausibility: it is a stretch to imagine fission cases, and imagining doppelgangers and the like is a step beyond this. Nevertheless, it seems to me that, while the addition of the doppelganger or the substitution of Keanu for the phone call do not affect the (narrowly-construed) causal relations between earlier and later objects, they *are* relevant to the question of whether the earlier object survives as the later. If I am correct, then we need a supervenience principle which permits causal influences on a region to make a difference to the persistence facts about that region, but does not permit the causal effects of a region to make a difference: the presence of extra offshoots can't undermine persistence, but the presence of extra sources can.

Compare my straightforward transfer with my merger with my doppelganger. Whether the later state casually depends on someone other than me makes a difference as to whether it is *my* later state. And compare my disrupted transfer with my Keanu-merger: again, the causes of the later state seem relevant. Where I fuse, with the doppelganger or Keanu, the existence

of a later person stage is not counterfactually dependent upon my earlier state: in my absence, either my doppelganger or Keanu would have sufficed. In contrast, when I am straightforwardly transferred, the existence of the later stage is counterfactually dependent upon my earlier state: in my absence, the telephone call would not have resulted in the existence of a person.<sup>20</sup> These differences in counterfactual dependence make a difference to what we should say about these cases: I can perish in fusion even though the fusion product is indiscernible from someone who would have been me in the absence of fusion. This marks a sharp difference from fission, where anti-CC intuitions suggest that I do not perish by having an extra offshoot.

To formulate a satisfactory supervenience principle we must build this causal asymmetry into the supervenience base. So: (for the target sortal F) two regions cannot differ in whether they at least partly contain an F without differing either intrinsically, or in the external causal influences which bear on the regions.

## V

Well, maybe this supervenience principle rules out closest-continuer theories of persistence, whilst respecting the asymmetry between fission and fusion, and without ruling out more acceptable ways in which facts about persistence may be externally influenced. If this version isn't quite right, maybe some close relative does the job. But it's not clear that our grounds for such a principle are any stronger than those for our initial intuitive rejection of closest continuer accounts. The principle does not seem to be of wider scope than that initial intuition, and it doesn't gain support from a wider theoretical context. But nor is it more transparently true than the initial intuition that whether or not an object persists cannot depend upon the (non)existence of a rival in the way that CC accounts suppose. We haven't done much more than reformulate the *only a and b* rule.

The pessimistic prognosis is that the *only a and b* objection to closest-continuer accounts results from confusion about contingent identity. Another possibility is that this is a stopping

---

<sup>20</sup> I have resorted to talk of person stages here in an attempt to avoid begging questions of identity without making the text unreadable. But the reference to stages is eliminable.

point: we just have the intuition that persistence cannot work that way, and no justification is either required or available. A more interesting possibility is that some general or more fundamental principle lies behind and sustains these objections, and the associated supervenience principle. What could this principle be?

One clue is that closest continuer accounts entail unusual counterfactual correlations between distinct objects. Compare a situation where  $f$  straightforwardly persists with a situation in which  $f$  divides symmetrically into  $g$  and  $h$ , and thus perishes (according to CC accounts). If  $h$  had not existed, neither would  $g$ :  $f$  would have continued to exist. Similarly, if  $g$  had not existed, neither would  $h$ . The two fission products,  $g$  and  $h$ , depend counterfactually upon one another for their existence. We may suppose that  $g$  and  $h$  are causally isolated, so they depend upon each other without mutually sustaining one another causally. Similarly, in the first situation  $f$  continues to exist only because  $g$  and  $h$  do not exist, but this is not because either  $g$  or  $h$  would have causally prevented  $f$ 's continuing existence.

In contrast, when fusion occurs, the counterfactual correlations entailed by closest-continuer accounts are causally explicable. Compare a situation in which  $j$  straightforwardly persists with a situation in which  $j$  fuses with  $k$ , both perish, and  $l$  emerges. Unlike fission products,  $j$  and  $k$  do not counterfactually depend upon one another for their existence. The existence of the fusion product,  $l$ , does depend upon that of both  $j$  and  $k$ , but this is unsurprising: after all  $j$  and  $k$  are the causal sources of  $l$ . Finally, when fusion does not occur,  $j$  continues to exist only because  $k$  does not exist. But again this is unsurprising: if  $k$  had existed, it would have interacted causally with  $j$ .

So closest continuer accounts seem unacceptable where they entail noncausal counterfactual correlations (fission), but acceptable where they do not have this consequence (fusion).

What's wrong with noncausal counterfactual correlations? It's not that beginnings and endings must always have causes, for such events could be indeterministic. It's not the absence of cause which is puzzling, but the correlation without the cause. CC accounts are

objectionable because they posit unexplained correlations; this is a more general theoretical vice.<sup>21</sup>

*Why* is it that *f* perishes when the rival branch is created? Our supposed reasons to believe this are as follows: given that there's only one object before the fission, and two distinct objects afterwards, and given that we cannot identify one fission product rather than the other with *f*, the laws of logic entail that *f* stops existing at fission. While the premises may support the conclusion, they do not seem to provide the right sort of explanation of *f*'s demise. We have seen that no causal explanation is possible. Nor does it seem plausible that this is a brute, inexplicable fact: *f*'s demise in fission is supposedly part of a modally robust and predictable pattern.

Sometimes we can explain why A is noncausally correlated with B by pointing out that one is somehow constitutive of the other. For example, if x is a part of y, or if x is identical to y, this explains all sorts of correlations between them. Is such a constitutive explanation available here? Can it be that in some sense, what it is for *h* to come into existence is for *f* to stop existing? According to CC accounts, if the first happens, so does the second. But is there some constitutive connection between them which explains their correlation? Showing that there is no such connection is the key to showing why CC accounts are unacceptable.

## VI

(How) can the existence or nonexistence of some object be constitutive of the continuing existence or failure to exist of some distinct object, without having a causal impact upon it? That's to say, (how) can there be objects which obey 'closest continuer' persistence conditions? It turns out that answers to these questions turn on broader questions in ontology. To explore these, I will adopt Eli Hirsch's very useful framework.<sup>22</sup> Hirsch points out that some regions of space-time seem to be occupied by ordinary 'natural' objects, whereas others seem gerrymandered; objects which occupied these regions would be 'unnatural'. Using this

---

<sup>21</sup> Didn't I earlier argue that we shouldn't criticise CC accounts for their claims about trans-world identity? Criticism based on these peculiar counterfactual correlations is not a criticism based directly upon transworld identity. Rather, it is criticism based on claims about connections between actual objects.

<sup>22</sup> Hirsch 1993.

*prima facie* division, Hirsch distinguishes three ontological views. Egalitarians argue that both natural and unnatural objects exist, and that the distinction between them is merely in the eye of the beholder. Elitists argue that both natural and unnatural objects exist, but that there is an objective difference between them, which may or may not be analysable. Finally, ontological inegalitarians deny that unnatural objects exist: gerrymandered regions of space-time do not exactly contain anything at all.<sup>23</sup>

As described here, these positions are under-specified, not least because there are different views about where the natural-unnatural boundary falls, and about whether this is a sharp boundary. Nevertheless, they provide a useful framework for our present investigation because egalitarians, elitists and ontological inegalitarians should take quite different attitudes to CC accounts of persistence. Or so I shall argue.

What makes an object natural or unnatural? (An object can be ‘natural’ in this sense even if it is a Barbie doll.) Hirsch focuses on features such as contrast with the environment, causal integrity, and so on: let’s call any object which qualifies as natural on these counts ‘actually natural’. Hirsch also mentions modal naturalness and unnaturalness.<sup>24</sup> An object is modally natural if it has natural ‘transworld identity’ conditions. To conceive of a modally unnatural object, take what you would normally think of as a non-rigid description and treat it as if it picked out the same object in a number of different worlds. For example, take ‘the first object that Kofi Annan thought about on 4th November 2003’. The object of this description is a bowl of cornflakes in the actual world, a Barbie in a nearby world, Keanu Reeves in another nearby world, and so on. If there is such an object, it is modally unnatural.<sup>25</sup>

Now, suppose there were objects which obeyed closest-continuer persistence conditions; call these ‘CC objects’. Would CC objects be either actually or modally unnatural? If so, how

---

<sup>23</sup> The threefold division cuts across the endurance-perdurance distinction: whilst endurantists are typically ontological inegalitarians, and perdurantists are not, these are not necessary connections.

<sup>24</sup> In addition, Hirsch raises questions about sortals: these involve both actual and modal naturalness.

<sup>25</sup> Counterpart theorists treat a modally unnatural ‘object’ as an object considered together with a unnatural counterpart relation; according to transworld-identity theorists, modally unnatural objects are unnatural in their own right. Where counterpart theorists see one object and many counterpart relations, identity theorists see many (co-located) objects. I will use the language of counterpart theory, but I intend my conclusions to be neutral between these two approaches.

damaging is this to closest-continuer accounts of persistence? I will consider these questions from the three ontological perspectives in turn.

## VII

According to egalitarians, the world contains a very large number of objects, each of which stands in a very large number of counterpart relations to other possible objects, and although we are more interested in some of those objects and relations than in others, this marks no deep, objective difference between them. There is carving, but there are no joints at which to carve. Egalitarians face no obstacle to belief in objects which obey closest-continuer persistence conditions. That's to say, given that a 'fission' event occurs, the egalitarian accepts that there is an object which ceases to exist at that moment, but which would have continued to exist otherwise. And where fission does not occur, there is nevertheless an object which would have ceased to exist had it divided into two equally good candidates. In addition, there are objects which persist through fission, and objects which cease to exist even in the absence of fission. Many of these objects coincide, at least temporarily: this may be understood either as the sharing of temporal parts or as 'constitution without identity'.

For egalitarians, there is no further question about which of these objects are objectively natural. And there is no sense in trying to explain why certain objects are noncausally correlated with other objects: given the egalitarian ontology, there are objects which fit pretty much any pattern of correlation you can imagine. The only remaining question is whether objects falling into ordinary categories like 'person', 'ship' and 'mountain' obey closest-continuer persistence conditions

Egalitarians accept that, before fission occurs, there are (at least) three coincident objects present, one of which perishes in fission, and two of which survive but become spatially separated. So egalitarians who oppose a closest-continuer account of F-persistence must argue that the surviving objects are better candidate Fs than is the one which perishes; this argument must be grounded in facts about us and our classificatory habits. Whether this can

be done depends upon the details of egalitarian theories of meaning and concept formation; some common strategies are unavailable to egalitarians.<sup>26</sup>

## VIII

Like egalitarians, elitists believe in many objects, but unlike egalitarians they believe that some objects are objectively more natural than others. Like egalitarians, standard elitists accept that there are objects which obey closest-continuer persistence conditions, and they may likewise wonder whether such objects fall under sortals like 'person'. But unlike egalitarians, elitists may approach this question by asking whether such objects are objectively natural.

An object which perishes in fission seems thereby to qualify as actually unnatural: it ceases to exist without the usual trappings of qualitative discontinuity, contrast with the environment, or disruption of internal causal processes. The challenge for elitist opponents of CC accounts is to justify this judgment of unnaturalness without simply reiterating the claim that CC accounts of the persistence of ordinary (natural) objects are unacceptable.

What about objects which have CC persistence conditions, but do not encounter fission? They are actually natural, but they would have ceased to exist in fission, so could have been unnatural. Are they therefore modally unnatural, with unnatural counterpart relations? We might think that no natural counterpart relation could relate a natural to an unnatural object, but this seems wrong. 'Same portion of matter' is a relatively natural counterpart relation, yet can relate a well-demarked chunk of matter to one which is not well-demarked.

Perhaps CC objects are modally unnatural because they involve counterpart relations which do not track genuine similarities. Consider Actual Kant and consider a possible Curtailed Kant, who expires in a pre-Critical fission event. Are they genuinely similar? Well, Curtailed Kant's life is an intrinsic duplicate of the early life of Actual Kant. But, let's suppose, Curtailed Kant coexists with Duplicate Kant, who survives fission and is an intrinsic

---

<sup>26</sup> Egalitarianism may be a way of at least modelling the view that, in some sense, we could have had other, equally good concepts of 'person'. See Merricks 2001 for trenchant criticism of this view.

duplicate of Actual Kant. The only similarity between Actual Kant and Curtailed Kant which does not hold between Actual Kant and Duplicate Kant is the following: in neither Actual Kant nor Curtailed Kant is an earlier stage psychologically and causally continuous with a later person-stage which coexists with later stages of the object. Approximately speaking, they both lack rivals. But this similarity is both extrinsic and negative: Actual Kant is more closely similar to Duplicate Kant than to Curtailed Kant, so the pairing of Actual Kant with a counterpart relation which selects Curtailed Kant is a ‘modally unnatural object’.

CC objects seem modally unnatural. Could they nevertheless be ships, people or mountains? On one understanding, unnatural objects cannot fall under genuine sortals, but this just raises the question of whether *person*, for example, is a genuine sortal. Instead, we should consider directly whether ‘is a person’ applies to unnatural objects. For elitists, as with egalitarians, debate about closest-continuer accounts seems to come down to debate about reference and predication. Both natural and unnatural objects are available as referents for our terms, and we need some story about which is selected. Elitists have the extra theoretical resource of an objective natural-unnatural distinction.

Do elitists even need an argument against closest continuer accounts? Lewis-style perdurantists who combine unrestricted composition with a counterpart theory of modality can easily adopt elitism.<sup>27</sup> They claim that there are (at least) two F-objects before a fission event, which then become spatially separated: they can thus maintain a continuity account of F-persistence without being driven to a closest-continuer account. But if this counts in favour of this metaphysical view, it must be because coincident objects are more eligible referents than CC objects.

Recall that the CC view is committed to seemingly-inexplicable correlations between distinct objects. Doesn’t the present elitism have a similar commitment? *Why* is it that the number of coincident Fs present depends in part upon whether fission later occurs? Here, however, there is a constitutive explanation of this correlation: what it is for there to be coincident Fs is for some stages to be common parts of two Fs which are elsewhere noncoincident. Those who don’t accept the elitist ontology will not accept such an explanation, but it does seem to

succeed on its own terms, and to rely upon independently understood ideas about parts and wholes.<sup>28</sup> Elitists may thus have the resources to argue against CC accounts of the persistence of ordinary things.

## IX

Ontic inegalitarians do not believe in unnatural objects: they believe that the world is quite sparsely populated. For elitists and egalitarians, the question is whether CC or non-CC objects are better candidates to be people, ships, and so on, given that they all exist. But for ontic inegalitarians, the question about CC objects is not whether they are natural, but whether they exist.

Suppose that we have established that CC objects would be unnatural if they existed. Isn't this reason enough for the ontic inegalitarian to claim that they do not exist? I doubt it. There seems to be nothing in the very notion of an unnatural object which makes it impossible or even unlikely that there be such objects. The ontic inegalitarian needs a more complex argument, one which directly targets the correlations posited by CC accounts between the demise of an object and the existence of its offshoot, or between the existence of two 'branch' objects. There are two ways such an argument might run.

First, ontic inegalitarians might argue that substances cannot display the kinds of correlation displayed by CC objects, because substances are 'ontologically independent'. This point must, of course, be accompanied both by a defense of the substance ontology and by an argument that people, ships, and so on are indeed substances; I take it that neither Parfit nor Nagel, for example, would accept this claim.<sup>29</sup>

Second, ontic inegalitarians could consider a range of possible styles of explanation of the correlations, show that these fail, and challenge the CC advocate to come up with something better. Causal explanations are ruled out, as are explanations in terms of identity or shared

---

<sup>27</sup> Lewis 1976.

<sup>28</sup> It's possible for endurantists to be elitists. But do endurantist elitists have the explanatory resources which are available to perdurantist elitists? I will not pursue this question here.

<sup>29</sup> For contemporary work in this substance tradition, see Simons 1987, Fine 1995, Lowe 1998.

parthood, for it is evident that these objects (e.g. the two fission branches) are wholly distinct. What other types of explanation are there? Constitutive correlations may occur when objects or facts of one type are reducible to, nothing more than, or somehow produced by objects or facts of another type; consider claims of mind-dependence, or of the dependence of certain objects or facts upon God, or of the mental on the physical, the chemical on the physical, and so on. But in this case, the objects in question are of the same type, and their dependence is symmetric.

Might CC advocates claim that the correlations are simply brute? This would undermine the initial motivation for connecting persistence and continuity.<sup>30</sup> The initial idea was that we have a fund of common-sense or intuitive data about the kinds of events ordinary objects can and cannot survive; continuity of various kinds seems to be a common theme. For egalitarians and elitists, these intuitions are guides to which objects we refer to in ordinary talk. But for ontic inegalitarians, these intuitions, if they are a guide to anything, are a guide to what exists. Roughly, there is a persisting object if and only if there is continuity,

Now, the closest-continuer advocate recommends that we make an exception for fission cases: here, and only here, continuity is not enough. Given certain assumptions, and the logic of identity, we have good reason to believe that continuity comes apart from persistence in this case. But once we see that continuity does not always suffice for persistence, why think that this applies only during fission? A fission event is not one in which an object is forcibly prevented from continuing to exist, despite all the other circumstances being favorable to its continued existence. Fission events are distinctive merely in that we can *tell* (given certain assumptions) that continuity and persistence come apart during them. But once this is conceded, we seem to have no reason to believe that continuity guarantees persistence in other circumstances.

Here's an analogy. Suppose that *being A* is not usually observable, but that *being B* is. You conjecture that *being B* guarantees *being A*; some mechanism connects the two. But on the rare occasion when you can directly observe the presence or absence of *A*, *A* is absent while *B* is present. It might be reasonable to continue believing in an *A-B* mechanism if you thought

that observation interfered with this mechanism, or that the operation of the mechanism systematically made *A* unobservable, or that there was some common cause of the failure of the *A-B* mechanism and the observability of *A*'s absence. But without such special explanations, you ought to give up the belief that there is in general a mechanism which usually makes *Bs* into *As*.

Do CC advocates think that persistence and continuity are correlated because of a 'mechanism', which is somehow disrupted by fission? It's hard to understand what this mechanism could be, but this is part of a broader problem of understanding the project of providing necessary and sufficient conditions for persistence within an ontic inequality framework. For ontic inequality, once the continuity advocate admits that continuity under a sortal is only usually sufficient for persistence, we have no good reason to think that there is even usually such a connection. The shift from a straight continuity account of *F*-persistence to a closest-continuer account undermines the motivation for the whole class of continuity accounts. This criticism helps spell out the accusation I mentioned right at the beginning of this article: the move from straight continuity account to CC account is *ad hoc* not because it is dialectically reactive, but because it undermines the initial motivation for the position.

## X

My goal was to find some rationale for the widespread rejection of closest-continuer accounts of persistence. First I argued that, although it is the logic of identity which motivates such accounts, they are objectionable not for their claims about identity as such, but for their claims about persistence. Second, CC accounts are not objectionable for their claim that persistence is a contingent matter, but for their claim that it is contingent upon the existence of some distinct object. This seems to violate some principle about the supervenience of persistence facts upon local or intrinsic facts. But a supervenience principle which rules out CC accounts without ruling out too much needs to be subtle and complex, neither a candidate for direct intuition, nor clearly a consequence of any broader view.

---

<sup>30</sup> Johnston 1989 makes a point like this.

In particular, I found it necessary to distinguish between (unacceptable) CC accounts of fission and (more plausible) CC accounts of fusion. An object which encounters fusion genuinely differs from one which does not, while an object which is duplicated in ‘fission’ does not seem genuinely to be affected. Moreover, the CC account of fission involves noncausal counterfactual correlations between distinct individuals, whilst the CC account of fusion does not. I suggested that these correlations are the truly objectionable element of CC accounts: if the correlations are unexplained, this is a flaw in the theory, and reason to think that there are in fact no such correlations.

I distinguished between three broadly different approaches to ontology: the egalitarian, the elitist and the ontic inegalitarian. Objections to noncausal correlations between objects look very different from these three different perspectives. Neither egalitarians and elitists deny that objects display these correlations, for there are objects for every possible correlation. Instead, they must argue that objects displaying these correlations could not be the objects of ordinary thought and talk; this task is easier for elitists, who have the resources to argue both that objects displaying these correlations are objectively unnatural, and that our ordinary terms do not pick out objectively unnatural objects.

Ontic inegalitarians must devise a different argument. Their best option is to claim that, since fission or the existence of a rival can merely be a sign of the failure of the original to persist, and cannot explain that failure, we have no good reason to think that continuity without persistence does not occur elsewhere. That’s to say, the move to the CC account undercuts the rationale for adopting a continuity account in the first place, so it is an unstable position.

The ‘only a and b’ principle is widely accepted, and once we discard confusions regarding identity, contingency and supervenience, the objection to CC accounts seems to turn upon their commitment to unexplained correlations. But this objection must be developed in quite different ways according to background assumptions about ontology and, frequently, according to the particular sorts of thing in question. Considerations about the naturalness of persons and about our epistemic access to facts about persons may differ radically from analogous considerations about ships, mountains or even cats.

## References

- Armstrong, David. 1980. Identity Through Time. In *Time and Cause*, edited by Peter van Inwagen, 67-78. Dordrecht: D. Reidel.
- Callender, Craig. 2001. Humean Supervenience and Rotating Matter. *Mind* 110: 25-44.
- Fine, Kit. 1995. Ontological Dependence. *Proceedings of the Aristotelian Society* 95: 269-90
- Gallois, Andre. 1998. *Occasions of Identity*. Oxford: Clarendon Press.
- Garrett, Brian. 1987. A Further Reply to Noonan. *Analysis* 47: 204-07.
- 1988. Identity and Extrinsicness. *Mind* 97: 105-09.
- 1990. Personal Identity and Extrinsicness. *Philosophical Studies* 59: 177-94.
- Hawley, Katherine. 1999. Persistence and Non-Supervenient Relations. *Mind* 108: 53-67.
- 2001. *How Things Persist*. Oxford: Oxford University Press.
- Heller, Mark. 1987. The Best Candidate Approach to Diachronic Identity. *Australasian Journal of Philosophy* 65: 434-51.
- 1990. *The Ontology of Physical Objects*. Cambridge: Cambridge University Press.
- Hirsch, Eli. 1993. *Dividing Reality*. New York: Oxford University Press.
- Johnston, Mark. 1987. Human Beings. *Journal of Philosophy* 84: 59-83.
- 1989. Fission and the Facts. *Philosophical Perspectives* 3: 369-97.
- Lewis, David. 1976. Survival and Identity. In *The Identities of Persons*, edited by Amelie O. Rorty, 17-40. University of California Press.
- 1983. New Work for a Theory of Universals. *Australasian Journal of Philosophy* 61: 343-77.
- 1986. *On the Plurality of Worlds*. Oxford: Basil Blackwell.
- 1999. Zimmerman and the Spinning Sphere. *Australasian Journal of Philosophy*, 77: 209-12.
- Lowe, E.J. 1989. *Kinds of Being*. Oxford: Basil Blackwell.
- 1998. *The Possibility of Metaphysics*. Oxford: Oxford University Press.
- Mackie, Penelope. 1987. Essence, Origin and Bare Identity. *Mind* 96: 173-201.
- Merricks, Trenton. 1998. There are no Criteria of Identity Over Time. *Nous* 32: 106-24.
- 2001. Realism about Personal Identity Over Time. *Philosophical Perspectives* 15: 173-87.
- Noonan, Harold. 1985a. The Closest Continuer Theory of Identity. *Inquiry* 28: 195-229.

- 1985b. Wiggins, Artifact Identity and 'Best Candidate' Theories. *Analysis* 45: 4-8.
- 1985c. The Only X and Y Principle. *Analysis* 45: 79-83.
- Nozick, Robert. 1981. *Philosophical Explanations*. Cambridge, MA: Harvard University Press.
- Parfit, Derek. 1971. Personal Identity. *Philosophical Review* 80: 3-27.
- 1984. *Reasons and Persons*. Oxford: Oxford University Press.
- Salmon, Nathan U. 1982. *Reference and Essence*. Oxford: Basil Blackwell.
- Shoemaker, Sydney. 1984. Personal Identity: A Materialist's Accounts. In *Personal Identity*, by Sydney Shoemaker and Richard Swinburne, Oxford: Basil Blackwell.
- Sidelle, Alan. 2000. Finding an Intrinsic Account of Identity: What is the Source of Duplication Cases? *Philosophy and Phenomenological Research* 61: 415-30.
- Sider, Theodore. 1996. All the World's a Stage. *Australasian Journal of Philosophy* 74: 433-53.
- 1999. Global Supervenience and Identity Across Times and Worlds. *Philosophy and Phenomenological Research* 59: 913-37
- 2001. *Four-Dimensionalism*. New York: Oxford University Press.
- Simons, Peter. 1987. *Parts*. Oxford: Oxford University Press.
- Wasserman, Ryan. forthcoming. Humean Supervenience and Personal Identity. *Philosophical Quarterly*.
- Wiggins, David. 1980. *Sameness and Substance*. Oxford: Basil Blackwell.
- 2001. *Sameness and Substance Renewed*. Cambridge: Cambridge University Press.
- Williams, Bernard. 1960. Bodily Continuity and Personal Identity. *Analysis* 21.
- 1970. The Self and the Future. *Philosophical Review* 79: 161-80.
- Zimmerman, Dean W. 1998. Temporal Parts and Supervenient Causation: The Incompatibility of Two Humean Doctrines. *Australasian Journal of Philosophy*, 76: 265-88.
- 1999. The Compatibility of Materialism and Survival: The 'Falling Elevator' Model. *Faith and Philosophy* 16: 194-212.