1 Identity Criteria

1.1 What Are Identity Criteria and Why Should We Care about Them?

Here are some examples of identity criteria:

- Object x is identical with object y just in case x and y have all the same properties.
- Sets A and B are identical if and only if A and B share all and only the same members.
- Events x and y are identical if and only if they are comprised of the same subjects, properties (and relations), and time intervals.
- Person x at t is identical with person y at t′ if and only if y at t′ is psychologically continuous with x at t′.

Identity criteria are powerful tools for the metaphysician. Equipped with identity criteria, theories gain predictive power. Suppose we had a comprehensive identity criterion that tells us when a person at one spatiotemporal location is identical with a person at another spatiotemporal location. Such a criterion will tell us whether individual identity and distinctness claims – such as "Mark Twain in 1860 = Samuel Clemens in 1875" and "Mark Twain in 1860 ≠ Harriet Tubman in 1854" – are true or not. We could also use it to answer questions like these:

- Will Cora survive the transition from being a rebellious, skateboarding seventeen-year-old to being a conservative forty-five-year-old investment banker?
- Will Harry survive a trip through a teletransporter that disassembles and then reassembles his physical matter?

An identity criterion for personal identity would tell us which changes we would survive and which changes would kill us. Likewise, identity criteria for events, facts, properties, material objects, actions, and objects in general would determine whether entities in those categories are identical or distinct under various circumstances.

Identity criteria can also help us shed light on our inferences involving identity claims. For instance, we know that the American author Mark Twain has an acerbic sense of humor. When we learn that Mark Twain is identical with Samuel Clemens ("Mark Twain" is Clemens' pen name), we attribute that acerbic sense of humor to Samuel Clemens. Although that inference seems obvious, we may wonder what principle(s) license it. Identity criteria, such as Leibniz's Law, can answer that question. Leibniz's Law, named after Gottfried Wilhelm von Leibniz, states – roughly – that individuals are identical if and only

if they share their properties. If Mark Twain has the property of *possessing an acerbic sense of humor*, then Leibniz's Law tells us that if Mark Twain = Samuel Clemens, then Samuel Clemens has that same property.

Although an identity criterion like Leibniz's Law helps license plausible inferences, it can also be a source of mystery. While it is controversial whether humans can survive trips through teletransporters, most of us agree that a person can survive the loss of a single strand of hair. But suppose that Amelia loses a single strand of hair at 12 p.m. on Tuesday. Amelia at 11:59 a.m. has 100,000 strands of hair while Amelia at 12:01 p.m. has 99,999 strands of hair. Prima facie, 11:59 Amelia possesses a property (*100,000 strands of hair*) that 12:01 Amelia lacks. Leibniz's Law seems to tell us that 11:59 Amelia is distinct from 12:01 Amelia, contrary to our initial judgment that Amelia survives losing a strand of hair.

This Element has two objectives: to discuss formulations of identity criteria and to take a closer look at Leibniz's Law. The first section concerns the general form of identity criteria. I address varieties of identity criteria present in the metaphysics literature and compare them. After providing an overview of varieties of identity criteria, I turn to a focused discussion of Leibniz's Law. Leibniz's Law is a conjunction of two principles of object individuation, the indiscernibility of identicals and the identity of indiscernibles. The second section concerns the better-regarded half of Leibniz's Law, the indiscernibility of identicals. The indiscernibility of identicals states that *if* objects x and y are identical, then x and y share their features. This principle seems so obvious to some that it even strikes them as akin to a logical truth. After all, if x and y are numerically one and the same object, that object must have all the same properties as itself. How can an object have different properties from the ones it itself has? Nevertheless, as we witnessed in the previous paragraph, there are challenges to the indiscernibility of identicals, and I will explain how they arise. In the third section, I turn to the more controversial half of Leibniz's Law, the identity of indiscernibles. The identity of indiscernibles states that if objects x and y share all the same properties, they are identical. This principle, depending on how it is interpreted, is less obvious. Why could there not be distinct objects – spheres, eggs, subatomic particles, or what have you – that share their features? We can imagine two eggs that have the same shape, size, color, and density along with the rest of their characteristics, can we not? I will explore alleged counterexamples to the identity of indiscernibles as well as alternative principles of object individuation that may serve as attractive competitors to the identity of indiscernibles. Finally, I will consider the possibility that there are no adequate criteria for object identity to be found.

1.2 Varieties of Identity Criteria

1.2.1 Material and the Modal Identity Criteria

At the most basic level, an identity criterion tells us when entities are numerically identical or distinct. There are many ways to formulate identity criteria, and I will examine some of their variations. First, I focus on the modal strength of identity criteria. Let us consider Leibniz's Law, the identity criterion that will be the subject of focus in the following sections. Here is one standard formulation of an identity criterion using Leibniz's Law:

Material Leibniz's Law: $\forall x \forall y (x = y \equiv \forall P (Px \equiv Py))$

This criterion provides necessary and sufficient conditions for the identity of individuals x and y: x and y are identical just in case they share all their properties. The x- and y-quantifiers range over objects, while the P-quantifier ranges over monadic properties. We will plausibly need to restrict the P-quantifier so that it does not range over *all* monadic properties. I will leave such restrictions to the next two sections.

Material Leibniz's Law is an *extensional* or "material"[1] identity criterion: it only provides a criterion for identity and distinctness for entities in the actual world; hence, we call it "Material Leibniz's Law." We can contrast this with a modalized version of Leibniz's Law, which is an instance of a Modal Identity Criterion:

Modal Leibniz's Law: $\Box \forall x \forall y (x = y \equiv \forall P (Px \equiv Py))$

Modal Leibniz's Law states that, necessarily, x is identical with y if and only if x and y share all their properties. Material identity criteria lack this modal strength. Material Leibniz's Law holds when all distinct objects differ with respect to at least one property at the actual world, and all identical objects do not differ with respect to any of their properties at the actual world. Certain popular counterexamples to Modal Leibniz's Law will not impact Material Leibniz's Law as long as those counterexamples do not describe actual states of affairs.

One popular counterexample to a version of Leibniz's Law is Max Black's sphere world (Black 1952). Black imagines a universe containing only two spheres with the same shape, mass, color, and all other physical characteristics. Let us suppose that the spheres reside five meters apart. The spheres are called "Castor" and "Pollux." Presumably, Castor and Pollux share all their qualitative properties (more on this notion in Section 3) even though they are distinct. One

[1] See Fine (2016).

may attempt to distinguish the spheres on the basis of their locational properties, but Black doubts this would work. If we maintain that Castor is distinct from Pollux because Castor has the property *located at spacetime region a* and Pollux has the property *located at spacetime region b*, then this only pushes the problem back: how can we distinguish between spacetime region a and spacetime region b? Do these regions not share all *their* properties?

Black's case challenges Modal Leibniz's Law, which states that, necessarily, if x and y are distinct, they differ with respect to their properties. If this is a metaphysically possible scenario in which distinct objects share all their properties, Modal Leibniz's Law must be false. While such a case may present a counterexample to Modal Leibniz's Law, it will not present a counterexample to Material Leibniz's Law because the actual world is not one that contains only two spheres floating in empty space.

It is somewhat difficult to find pressing counterexamples to Material Leibniz's Law. Because the universe does not appear to involve a symmetrical distribution of matter, we should often be able to distinguish actual objects by the different relational properties they have. Even if we have two duplicate spheres, Actual Castor and Actual Pollux in the actual world, we could find relational properties to distinguish them. For instance, suppose Actual Castor and Actual Pollux materialized in the state of Alabama. Figure 1 shows a map of where they materialized.

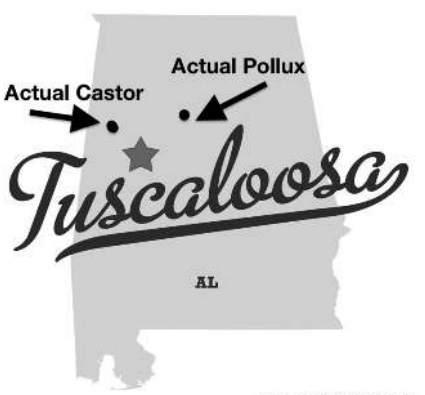

Figure 1 Map of the locations of Actual Castor and Actual Pollux in Alabama. The map is a modified version of the map of Tuscaloosa, Alabama, found on TownmapsUSA.com. I have added the dots, names, and arrows used to represent Actual Castor and Actual Pollux. License for use: https://creativecommons.org/licenses/by-nc/3.0/

Given their locations, we can see that Actual Castor and Actual Pollux differ in their distance from Tuscaloosa. Actual Castor and Actual Pollux differ with respect to the following properties: Actual Castor is 50 miles from Tuscaloosa. Actual Pollux is not. Actual Pollux is 100 miles from Tuscaloosa. So even though Actual Castor and Actual Pollux may be indistinguishable with respect to their mass, temperature, shape, color, and so on, we can distinguish them by their relational properties. Furthermore, there are many related properties that will distinguish Actual Castor and Actual Pollux. It just so happens that the Alabama Museum of Natural History (located in Tuscaloosa) contains an intact skull of an American Mastodon. Actual Castor has the property of *being 50 miles from the skull of a mastodon*, while Actual Pollux lacks this property and instead has the property *being 100 miles from the skull of a mastodon*. The existence of Actual Castor and Actual Pollux does not pose a counterexample to Material Leibniz's Law.

This is not to say that we cannot find counterexamples to Material Leibniz's Law (see Section 3 for further discussion of this), but potential counterexamples to Modal Leibniz's Law will not immediately serve as counterexamples to material versions. I have been discussing Leibniz's Law specifically, but this lesson should apply to formulations of identity criteria in general. Material versions of identity criteria are weaker than their necessitated counterparts, and it is easier to find potential counterexamples to the modal versions than it is to the material ones.

Greater resistance to counterexamples provides a prima facie reason to favor material identity criteria over modal identity criteria. But there are costs to embracing material identity criteria over modal ones. Let us return to the case of Leibniz's Law. If we adopt Material Leibniz's Law instead of Modal Leibniz's Law, we face the question of why this identity criterion is only contingently true.

Some philosophers have proposed that at least certain identity claims of the form "x = y" are necessarily true if true. I am thinking specifically of philosophers such as Saul Kripke (1980), who believe that when identity claims include proper names flanking the two sides of the identity predicate the claim is necessarily true. For Kripke, proper names (like "Mark Twain") refer to the same individual in every possible world in which the individual exists. They are "rigid designators" in his terminology. If "Mark Twain" and "Samuel Clemens" refer to the same individual in the actual world, they do so in every possible world. Accordingly, the identity claim "Mark Twain = Samuel Clemens" is true in every possible world if it is true in the actual world. It would be surprising for certain identity and distinctness facts involving individuals to obtain necessarily even though the identity criteria for such individuals do not obtain necessarily.

Relatedly, whether we will be satisfied with material identity criteria depends on what we want to use the identity criteria for. Often, we ask questions like: Were I to lose half of my brain in an accident would I be numerically one and the same person? Were this set to contain an extra member, would it be the same set? And, if some event had taken place earlier would it have been the same event? These questions concern counterfactual circumstances. If our identity criteria are to be useful in answering these questions, we would expect them to be *modally stable* – true in at least the other possible worlds we are concerned with when asking these questions. For these reasons, we may search for modal identity criteria rather than material ones. These are not conclusive reasons to adopt a modal identity criterion over a material one. My intention is merely to highlight some issues that arise when opting for one type of identity criterion over the other.

Modal identity criteria have a necessity operator appearing with wide scope, but there are varieties of necessity operators. Modal identity criteria can come in different strengths, depending on what notion of necessity we deploy. We may maintain that the biconditional holds with metaphysical necessity or rather with physical necessity. In the former case, the identity criterion is supposed to hold in every metaphysically possible world. In the latter case, the biconditional would hold in every possible world that is consistent with the actual laws of nature.[2] In metaphysical discussions of identity criteria, philosophers often have the metaphysical rather than the physical or nomological necessity operator in mind.

1.2.2 Explanatory Identity Criteria

We can also distinguish between *explanatory* and *nonexplanatory* identity criteria. Explanatory identity criteria do not provide necessary and sufficient conditions for identity and distinctness – at least not directly. Instead, they tell us *in virtue of what* identity and distinctness facts hold. In recent years, explanatory identity criteria have taken the form of grounding criteria for identity and distinctness.[3] We can understand the *in virtue of* relation in terms of ground. We can convert the previous modal and material identity criteria for set identity and Leibniz's Law to explanatory identity criteria as follows:

[2] I will speak of different metaphysical possibilities as different metaphysically possible worlds, but nothing I say should hinge on adopting a possible worlds framework as opposed to another framework about modality.

[3] See Burgess (2012), Fine (2016), Shumener (2020a; 2020b), and Wörner (2021).

Explanatory Set Identity:

> If set x = set y then x = y is fully grounded in the fact ∀s (s is a member of x ≡ s is a member of y)
> If set x ≠ set y then x ≠ y is fully grounded in the fact ∃s ((s is a member of x ∨ s is a member of y) & ~(s is a member of x & s is a member of y)).

Explanatory Leibniz's Law:[4]

> If x = y then x = y is fully grounded in the fact ∀P(Px ≡ Py)
> If x ≠ y then x ≠ y is fully grounded in the fact ∃P((Px ∨ Py) & ~ (Px & Py)).

What is ground? The notion of ground has been popularized over the past decade by many philosophers.[5] Metaphysicians typically understand ground as either a relation (holding among facts) or a sentential operator. When x grounds y, y holds in virtue of x. Ground is supposed to be either a type of metaphysical explanation or a relation that backs metaphysical explanation, depending on which grounding theorists one consults.[6] When a fact is ungrounded, it is not grounded by any further facts. I take ungrounded facts to be metaphysically fundamental.

There are many contexts in which we want to claim that certain facts hold in virtue of other facts. For example, we want to determine whether:

- normative facts obtain in virtue of descriptive ones (e.g., does "Action x is morally required" hold in virtue of "Action x maximizes happiness"?)
- mental facts obtain in virtue of physical facts (e.g., does "s is in pain" hold in virtue of "s's c-fibers are firing"?)
- determinable facts obtain in virtue of determinate facts (e.g., does "Annie's shirt is red" hold in virtue of "Annie's shirt is scarlet"?)
- disjunctive facts obtain in virtue of their disjuncts (e.g., does "Either Pittsburgh is in Kentucky or Pittsburgh is in Pennsylvania" hold in virtue of "Pittsburgh is in Pennsylvania"?).

Grounding theorists often propose that the same notion of *in virtue of* appears in these sentences, and we can formulate the corresponding claims in terms of ground.

- "Action x is morally required" is fully grounded in "Action x maximizes happiness."

[4] Technically, this is an explanatory version of only one half of Leibniz's Law, the identity of indiscernibles. We discuss this issue in Section 3.

[5] See Fine (2001; 2012), Schaffer (2009), Rosen (2010), Raven (2013; 2015), as well as many others.

[6] See Trogdon (2013), Thompson (2016), Maurin (2019), and Glazier (2020) for discussion of the relationship between ground and metaphysical explanation.

- "s is in pain" is fully grounded in "s's c-fibers are firing."
- "Annie's shirt is red" is fully grounded in "Annie's shirt is scarlet."
- "Either Pittsburgh is in Kentucky or Pittsburgh is in Pennsylvania" is fully grounded in "Pittsburgh is in Pennsylvania."

We may also accept general grounding principles concerning the facts in question: normative facts, in general, hold in virtue of descriptive facts in the sense that they are grounded in descriptive facts. If we agree with grounding theorists at this juncture, we may suppose that certain *in virtue of* claims hold for identity and distinctness facts as well, and that these claims should be understood in terms of ground. For instance, we may be inclined to accept:

- The fact that individual x is identical with individual y obtains in virtue of the fact that x and y share all their properties.
- Set x is identical with set y in virtue of the fact that x and y share all their members.
- Person x at t is identical with person y at t′ in virtue of the fact that y at t′ is psychologically continuous with x at t.

If these are statements of ground, then that suggests we will uphold explanatory identity criteria; perhaps there is a general principle (or set of general principles) telling us how identity and distinctness facts are grounded. In what follows, we appeal to a grounding relation that holds among facts.[7] A fact is grounded by another fact or a plurality of facts. We can also distinguish between full and partial ground. I take full ground to be a primitive notion. But intuitively, P fully grounds Q when Q holds in virtue of P and P is sufficient on its own to explain Q. P partially grounds Q when P on its own or together with further facts fully grounds Q. For instance, the fact that the scarf is scarlet *fully grounds* the fact that the scarf is red. The former fact suffices to ground the latter. But the fact that the sky is blue only partly grounds the conjunctive fact that the sky is blue and the grass is green. That conjunctive fact is fully grounded in the plurality of facts: the sky is blue, the grass is green.

Where P, Q, and R are facts, both full and partial ground should obey the following conditions:

Asymmetry: If P grounds Q, then Q does not ground P.

Transitivity: If P grounds Q and Q grounds R, then P grounds R.

[7] Some grounding theorists take ground to be a sentential operator rather than a relation holding between facts. See Fine (2012).

Irreflexivity: P does not ground P.

Grounding Necessitation: If P fully grounds Q then necessarily, if P then Q.

While these are popular constraints of ground, all of these constraints have been questioned or rejected by at least some philosophers working on ground.[8] Nevertheless, we will assume these constraints hold for now – this will allow us to get a better grasp of how we can deploy a mainstream conception of ground to understand identity criteria.

Now we have the basic understanding of the notion of ground – of the relation appealed to in explanatory identity criteria – I will highlight some ways in which explanatory identity criteria differ from modal and material identity criteria. Modal and material identity criteria do not entail corresponding explanatory identity criteria. For example, Material Leibniz's Law does not necessitate an explanatory relationship in either direction. We cannot conclude from Material Leibniz's Law that objects' having the same properties explains their identity. We also cannot infer from Material Leibniz's Law that the identity of objects x and y explains the fact that x and y share all their properties. Material Leibniz's Law only tells us that objects have the same properties *if and only if* they are identical. Modal Leibniz's Law does not entail corresponding explanatory identity criteria either. Even if it is metaphysically necessary that x = y if and only if x and y share all the same properties, we cannot, on that basis, conclude either (a) that x and y's sharing their properties explains x = y or (b) the fact that x = y explains x and y's sharing their properties.

It is compatible with modal and material identity criteria that no explanatory relationship holds whatsoever. The choice of whether to adopt explanatory identity criteria rather than (or in addition to) modal or material identity criteria will depend upon one's explanatory ambitions when advancing identity criteria. Why would one wish to defend explanatory identity criteria? If we think that, when providing identity criteria, we are stating *in virtue of what* identity and distinctness facts hold, and we take the *in virtue of* relation to be asymmetric, then this motivates an appeal to explanatory identity criteria.

We may be attempting to uncover this *in virtue of* relationship when we attend to various identity-related puzzles. When considering whether a person x entering a Star Trek–style transporter device is numerically identical with a person y who emerges from a Star Trek–style transporter device at a later time, we do not merely want to know whether x and y are numerically identical or

[8] For someone who questions asymmetry, see Koslicki (2015). For those who develop reflexive accounts of ground (often called "weak ground"), see Fine (2012) and deRosset (2013). For rejecting transitivity, see Schaffer (2012), and see Skiles (2015) for a rejection of grounding necessitation.

distinct, but also in virtue of what their identity or distinctness holds. Likewise, it would not be enough – some philosophers think – to establish that Castor and Pollux are distinct yet qualitatively indiscernible spheres, ones that share their mass, shape, density, temperature, color, and so on. We also want to know *why* Castor and Pollux are distinct, given that they are qualitatively indiscernible.

Not everyone is interested in this explanatory project when investigating identity criteria. There are some potential reasons to accept modal or material identity criteria and reject explanatory identity criteria. First, one may be skeptical of notions of metaphysical explanation or ground in general.[9] If we understand metaphysical explanation in terms of ground, and one is skeptical about ground, then one will deny that there are identity criteria that tell us how identity and distinctness facts are grounded.

But even if one accepts that some facts are metaphysically explained or grounded, one can deny that we need explanatory identity criteria. If identity and distinctness facts (at least some of them) are good candidates for fundamental facts, then it is not clear that we need explanatory identity criteria. One thought is that the identity relation is a primitive, logical notion, and perhaps some facts involving primitive logical relations need no explanation. Williamson (1990: 145) echoes the idea that identity facts do not need to be explained. He maintains that for any objects belonging to a kind F, we should not try to explain why they are identical. David Lewis (1986: 192–93) also claimed that "there is never any problem about what makes something identical to itself." Neither Lewis nor Williamson had notions of ground explicitly in mind in these passages, but if we are sympathetic to their claims, we may resist attempts to provide explanatory identity criteria.[10]

Another related reason to deny the existence of explanatory identity criteria is to claim that identity and distinctness facts are not explained; rather, they do the explaining.[11] For instance, identity and distinctness facts are not explained by objects sharing or differing in their properties; instead, the identity and distinctness of objects explain their sharing or differing in properties. In this case, the explanatory relation would point in the opposite direction from the direction it points in in the explanatory identity criteria listed earlier: identity and distinctness facts do not stand in need of explanation. They explain other facts. As such, we should deny the need for explanatory identity criteria here as well. If the identity and distinctness facts have this kind of explanatory power, we could treat such

[9] See Wilson (2014) and Koslicki (2015) for skepticism about ground.
[10] Also see Bueno (2014) for considerations in favor of taking identity to be fundamental.
[11] See Wilhelm (2021) for a fascinating discussion of the explanatory power of identity facts.