NOÛS 42:1 (2008) 1–16

# *Why Sense Cannot Be Made of Vague Identity*[1]

NICHOLAS J. J. SMITH
The University of Sydney

**0.** The idea that identity might sometimes be vague arises naturally in a range of cases. There have been a number of arguments to the conclusion that identity can never be vague, on pain of contradiction, and likewise a number of attempts to model vague identity in a clear and consistent way and thereby show that it is perfectly possible. I think it is fair to say that the majority of philosophers reject the idea of vague identity, but still find it puzzling. In this paper I present a new argument against vague identity, which is more fundamental than existing arguments, and I also try to explain why we find the idea of vague identity puzzling, in a way that will dispel the puzzlement. The two projects are closely linked: the conclusion of my argument is that it is not possible to make clear sense of vague identity, and explaining why this is so sheds light on the puzzlement we experience when we do try to make sense of it. Due to the special role that identity plays in the apparatus we use for making sense of any phenomenon whatsoever, vague identity becomes a 'no go' zone: if we can make clear sense of the possibility of a given relation's being vague, then that relation cannot be identity.

**1.** A smoke-stack belches out a large, roundish cloud of smoke. Air currents narrow the cloud in the middle—into the shape of a dog's bone—and then eventually divide it into two separate, roundish clouds of smoke, one on the left and one on the right. We start with one cloud of smoke; we finish up with two. At any point during this process, we can point to the left side of the cloud/the left cloud and call the cloud of which the bit we are pointing to is part 'Cloud L', and point to the right side of the cloud/the right cloud and call the cloud of which the bit we are pointing to is part 'Cloud R'. At points near the beginning of the process, it is clear that clouds L and R are identical (i.e., one and the same cloud). At points near the end of the process, it is clear that clouds L and R are distinct. At certain points in the middle of the process, it is unclear whether clouds L and R are identical.

A museum claims to have George Washington's axe on display, yet admits that the axe has had its handle replaced three times and its head replaced twice since Washington wielded it. Is the axe in the museum the very axe with which George Washington cut down the apple tree? On the one hand it seems not, as it contains no parts in common with the original axe. On the other hand, if it is not the same axe, then when was the old axe supplanted by the new (for surely an axe can survive the replacement of a broken handle, and likewise the replacement of a chipped head)? The Ship of Theseus is a more complex version of the same problem. Suppose that each timber on a wooden ship is replaced when it becomes rotten or damaged, until eventually every timber has been replaced. Is the resulting ship the original ship? In both these cases, we have an initial object and a final object, and there is continuity of form and function between them, but they have no physical constituents in common. The question whether the initial object and the final object are one and the same object then appears to have no clear answer.

These and other sorts of example prompt the thought that sometimes it is a vague matter—out there in the world—whether objects *a* and *b* are identical.[2] That is, it is a vague matter whether there are two objects *a* and *b*, or just one object picked out once as '*a*' and once as '*b*'. My question is, can we make clear sense of this idea?[3]

**2.** As so often in philosophy, before we can answer our question, we need to say a little about what it means. I claim that a necessary condition for *making clear sense* of a phenomenon is showing how the phenomenon may be modelled using standard set-theoretic tools. For example, Kripke did this for possible worlds when he presented a set-theoretic model theory for modal languages which employed such worlds, and Tarski and others did this for semantic notions such as truth and reference when they developed classical model theory. Many philosophers would demand more than this before they would agree that clear sense had been made of a phenomenon: for them, to make clear sense of something is to give a naturalistic (or perhaps physicalistic) account of it. From this point of view, Kripke did not make sense of possible worlds (to do that would be to see how they might be constructed out of naturalistically acceptable materials such as, say, sentences) and Tarski did not make sense of semantic notions (to do that would be to see how they might be reduced to naturalistically acceptable materials such as, say, causal chains). For the purposes of this paper, however, I do not need to enter into the question whether, once we have made sense of something set-theoretically, we need do anything more before we can say that we have made sense of it *simpliciter*. I only need to defend the claim that if we cannot model a phenomenon using set-theoretic tools, then we certainly cannot make sense of it in naturalistic terms, or indeed make clear sense of it at all, in any reasonable sense of 'make clear sense'. Why is this so? Because when we show that some phenomenon can be modelled using standard set theory, we show that the phenomenon is unobjectionable *from a purely logical point of view*. When we know that some statements can be modelled in set theory, we know that they *have a model* in the logical sense, and thus at the very least are consistent. If a phenomenon is not even *logically* coherent, then there certainly will not be a good *naturalistic* account of it.

This weak claim—that a *necessary* condition for making clear sense of a phenomenon is showing how the phenomenon may be modelled using standard set-theoretic tools—is all I need here: for my purpose is to show that we cannot model vague identity set-theoretically.

**3.** Identity has a special place in set-theoretic models: the identity facts are *built into* every such model. Each such model comes pre-equipped with a 'factory-installed' identity relation. Other relations can be added to the model, but the pre-installed identity relation cannot be removed.

Consider a couple of examples. The intuitive idea of *distance* is modelled in the theory of *metric spaces*. A metric space is a set of objects together with a mapping $d$ from pairs of objects to non-negative real numbers. The idea is that the number $d(a, b)$, to which the pair of objects $(a, b)$ is mapped, represents the distance between $a$ and $b$. There are various constraints which the mapping must meet. For example, for any objects $x$ and $y$ in the set, $d(x, y) = d(y, x)$, i.e., the distance between $x$ and $y$ must be the same as the distance between $y$ and $x$. Another constraint—the one of interest here—is that $d(x, y) = 0$ if and only if $x$ and $y$ are identical, that is if and only if $x$ and $y$ are one and the same object. Note what is going on here. We have a theory which posits a particular function—a distance function—and in specifying the properties of this function, we appeal to another relation—identity—about which we said absolutely nothing at the outset when we introduced our raw materials for a metric space. We said we have a set of objects, together with a function from pairs of such objects to non-negative real numbers; we did not add that we also have a relation (identity). We could not proceed in this way with any relation other than identity. We could not place some constraint on $d$ which referred to, say, the *larger than* relation amongst objects. If we wanted to do that, we would have to explicitly say that a metric space consists of a set of objects together with a relation on that set—the larger-than relation—and a function from pairs of members of the set to real numbers. But with identity we do not have to do this. We do not have to do this because as soon as we have a set of objects, we have the identity relation on that set: the relation holds between each object and itself, and between no object and any object other than itself.

Consider a second example: the model theory for first-order logic with identity. We have a standard first-order language, to which we add a new two-place predicate '=' (or '$I$', or whatever symbol we prefer). We now say that a model for our language consists of a set of objects (the domain), together with a mapping which assigns to each name an object in the domain, and to each relation symbol *except* '=' a relation on the domain. Note that at this stage we do not say that to specify a model, we also need a two-place relation of identity on the domain. And yet later, when we give the truth conditions for sentences of the form '$a = b$', we say that such a sentence is true just in case the referent of '$a$' is identical to the referent of '$b$'. This is quite different from how we proceed with other two-place predicates. '$R(a, b)$' is true just in case the referent of '$a$' and the referent of '$b$' stand in the relation referred to by '$R$'—whatever relation that happens to be in the model in question. But with identity we refer to a particular relation—identity—which is just there in the model in the first place, without our having to do anything to put it there. Once again, as soon as we have a bunch of objects—the domain—we immediately have the identity relation on that domain.

Someone might object that the identity relation is not really there on the domain to begin with: we put it there by a conventional definition; we stipulate that the extension of the identity predicate is the set of all ordered pairs $(x, x)$ for every object $x$ in the domain. But this stipulation only works if we assume (as we do) that both occurrences of '$x$' in '$(x, x)$' refer to *one and the same object*—in other words, that the thing that the first occurrence of '$x$' picks out is identical to the thing that the second occurrence of '$x$' picks out—and here again we have this primitive relation of identity, with which we assume the domain comes equipped. To spell this out in a little more detail: Suppose we have three logic textbooks. All of them say that on every interpretation $\mathfrak{M} = (D, I)$, consisting of a domain $D$ and an interpretation function $I$, $I$ assigns to the symbol '=' the relation $id_{\mathfrak{M}}$, the identity relation on $D$. One of them says nothing more about $id_{\mathfrak{M}}$: the author assumes that we know what the identity relation on $D$ is. The other two define $id_{\mathfrak{M}}$. One defines it thus:

$$id_{\mathfrak{M}} = \{(x, y) : x, y \in D \wedge x = y\}$$

Here the notion of real identity on the domain occurs on the right hand side. The other defines it thus:

$$id_{\mathfrak{M}} = \{(x, x) : x \in D\}$$

Here there is no undefined occurrence of '$id_{\mathfrak{M}}$' or '=' or 'identity' etc.—but nevertheless this notion plays an essential role in our understanding the definition—in particular, in our understanding '$(x, x)$'. We are to understand that we construct an ordered pair by taking an object in the domain in first place, and then taking *the very same object* (i.e., the one and only object that is identical to the first) in second place. If we do that for every object in the domain, we will have assembled the extension of '='. Thus the lack of any undefined occurrence of '$id_{\mathfrak{M}}$' or '=' or 'identity' etc. in the third textbook is an entirely superficial matter. It is with good reason that logicians would not regard the first two textbooks as sloppier than the third—as taking for granted something that the third spells out explicitly and accurately. Logicians would regard all three textbooks as saying the very same thing—for they all, explicitly or implicitly, appeal to a primitive identity relation with which any set of objects comes equipped.

I have said that as soon as we have a set of objects, we have the identity relation thereon. This is *not* to say that we have *criteria* of identity for the objects in the set. We may have no idea how to *determine* whether or not *a* and *b* are one and the same object, given some description (or other mode of presentation) of *a* and some description (or other mode of presentation) of *b*. It is just to say that we have all the identity *facts*. There are some objects in the set, and each one is identical to itself and not to any other object.

**4.** Now let us consider the possibility that the identity relation might be vague. My claim is that we cannot make sense of this idea, for the following simple reason. In order to make sense of the idea of a vague identity relation, we would need to model this relation set-theoretically. We would need, for example, to posit a fuzzy relation obeying some analogues of the principles governing identity (reflexivity, transitivity, symmetry, Leibniz's Law).[4] If we did not present a set-theoretic model of this sort, we certainly would not have made sense of vague identity. Yet if we do present a set-theoretic model of some vague relation obeying analogues of the principles governing identity, this modelled relation *will not be identity*, because as noted above, identity is already there, built into the background of every such set-theoretic model—and it is precise (the identity relation holds between each object and itself, and between no object and any object other than itself). In short, whereof we have not modelled set-theoretically, thereof we have not made sense; while any vague relation we do model set-theoretically will not be identity, for real identity will already be there, built in to the background of our model, and perfectly precise.

Some examples will help make this clear. Graham Priest [1998, 332–3] writes:

> Let us suppose that we are given a domain of objects, *D*, and that the objects come with a distance metric, *d*. Specifically, *d* is a non-negative real-valued function satisfying the conditions:
>
> $$d(x, x) = 0$$
>
> $$d(x, y) = d(y, x)$$
>
> $$d(x, z) \leq d(x, y) + d(y, z)$$
>
> where $x, y, z \in D$. If the distance between *x* and *y* is small under the metric, then they may not be *completely* identical, but they will be *almost* so. That is, the degree of their identity will be nearly 0. This suggests taking the metric itself as providing the degree of truth of an identity statement.

A problem arises for Priest at the outset, when he says "suppose that we are given a domain of objects". If we are to understand this in the usual way, then we have a set of objects; but as soon as we have a set of objects, we have the identity relation thereon—as discussed above, the identity facts are given automatically along with the set of objects. This means that when we introduce our metric on the set, and then take the metric as providing the degree of truth of statements of the form '$a = b$', we succeed only in making '=' mean something other than identity. We have not *made identity vague*. We have simply added some new structure to our domain, on top of the real identity relation with which it came pre-equipped. We can make whatever stipulations we like about how to interpret the symbol '=', but unless this symbol picks out the real identity relation on the domain, it does not mean identity; and we have done nothing to make this real identity relation vague—as noted, we have simply added extra structure to the domain.[5]

Later in his paper, Priest considers the objection that his notion of identity "is not true identity, but an *ersatz*" [Priest, 1998, 337]. Priest responds that this objection begs the question, and continues: "Plausibly, the fuzzy truth conditions do not define an *ersatz* identity relation, but the genuine thing for vague objects... The objector does not even have a right to *assume* that the domain $D$ is furnished with a relation of the kind required for classical identity" [Priest, 1998, 337]. This response gets right to the heart of the matter. In the ordinary understanding of set-theoretic models, we do assume that a set of objects comes equipped with an identity relation, which is always completely precise—and we have every right to assume this, because this is just how the standard view works (as illustrated in the previous section). Now if someone is seriously proposing to make sense of vague identity, she will indeed need to present a framework in which we can have some objects *without* automatically having the classical identity relation on these objects. *If* she makes such a proposal, and we then assume a classical identity relation, then we will indeed be begging the question, and making an unjustified assumption. But Priest has made no such proposal. His presentation proceeds in what looks like a perfectly standard way: we have a set of objects, a metric on the set is introduced, and a model theory for vague language is presented. So far so good: *assuming we are to understand this in the ordinary way*. But here's the dilemma: if we are to understand it in the ordinary way, then we *are* justified in assuming that the set comes pre-equipped with a classical identity relation; while if we are not to understand it in the ordinary way, then we need to be told how to understand it, or we do not understand it at all. Priest tells us not to assume a classical identity relation—*a fortiori*, not to understand his presentation in the ordinary way—but he does not give any other way for us to understand it.

To take another example, consider van Inwagen's [1988, 261–2] attempt to make sense of vague identity:

> A *universe* is a non-empty set of objects. A *pairing on* a universe is a (possibly empty) set of two-membered sets (pairs) of members of that universe. (These are to be "genuinely" two-membered: $\{x, x\}$ [$= \{x\}$] cannot be a member of a pairing.) If $x$ and $y$, $x \neq y$, are members of a pair (belonging to a certain pairing) they are said to be *paired* (in that pairing). [¶] A *model* consists of a universe, a pairing on that universe, and an assignment of one object in that universe to each individual constant... The objects with which an object is paired are to be thought of as the objects such that it is indefinite whether that object is identical with them.

Van Inwagen seems to be presenting a standard sort of set-theoretic model, and indeed makes free use of the relation of identity with which any ordinary set comes pre-equipped (both in specifying that pairs be genuinely two-membered, and when he says "If $x$ and $y$, $x \neq y$, are members of a pair...")—but this is then in tension with the later claim that "The objects with which an object is paired are to be thought of as the objects such that it is indefinite whether that object is identical with them." We have been explicitly told that in a pairing $\{x, y\}$, $x$ and $y$ must be non-identical. Now we are told that $x$ and $y$ are to be thought of as indefinitely identical. I cannot make sense of this. If we are working with a standard set-theoretic model, then $x$ and $y$ are simply non-identical; if we are not, then unless we are given some other way to understand the presentation, we do not understand it at all. The ordinary understanding of set-theoretic models rules out van Inwagen's proposed interpretation of his construction—yet no other way of understanding the construction is presented.

Contrast the two cases just considered (Priest and van Inwagen) with a case where a set-theoretic construction is given, and where the interpretation placed on the construction does *not* conflict with the ordinary understanding of it. Trillas and Valverde [1984, 233–4] present the following example: "Let $X$ be a screen where nebular clusters $F, G, \ldots$ of points appear (e.g., stains made by perpendicular spray-painting from a regular distance). We call these clusters "nebulae"." They then develop a theory in which (a) each nebula is (identified or associated with) a function from points in $X$ (i.e., points on the screen) to the set $\{0, \frac{1}{2}, 1\}$, so that for each nebula, a point is either in it (mapped to 1), not in it (mapped to 0), or this is indeterminate (mapped to $\frac{1}{2}$), and (b) two nebulae are said to *coincide* if they map all points to the same values, to be *distinguishable* if one maps some point to 1 and the other maps that point to 0, and to be *mixed up* if they neither coincide nor are distinguishable (i.e. there is no point which one maps to 1 and the other to 0, but there is a point which one maps to 1 or 0 and the other maps to $\frac{1}{2}$). This is the sort of case which some might be inclined to say involves vague identity: two nebulae which are mixed up in Trillas and Valverde's sense might be thought to be indefinitely identical. But we cannot think of matters thus if we understand the construction in the standard way. The two nebulae are non-identical (for they are different functions from $X$), but they do stand in *another* relation (being mixed-up). This relation cannot coherently be seen as indefinite identity. As Trillas and Valverde say [p.232]: "The indistinguishability... overlays identity—it does not supplant it."[6]

**5.** To sum up my argument in this paper:

1. To make clear sense of something one must (at least) model it set-theoretically.
2. Vague identity cannot be modelled set-theoretically.
3. Therefore we cannot make clear sense of vague identity.

Premise 2 has been illustrated in the foregoing examples. The task of understanding vague identity is in an entirely different league from the task of understanding vagueness in any other relation (e.g., 'loves', 'is near', 'is friends with', 'enjoys', ...). We can model the latter sort of vagueness within standard set theory.[7] Identity, however, is special. Vague identity is not something that can be modelled using set-theoretic apparatus, for classical identity is built into every model that we construct using that apparatus, and thus, any vague relation that we model using that apparatus is not identity.

What about Premise 1? I defended it at the outset, but it is time to say a little more—for I can imagine someone objecting:

> What system of set theory are you talking about? ZFC? There are things which cannot be modelled in ZFC (e.g., a universal set, a self-membered set) which arguably do make sense, and which can be modelled in other systems of set theory (e.g., NF, non-well-founded set theories). In general, whichever set theory you intend, being able to model a phenomenon within your chosen system of set theory cannot be taken as a mark of our being able to make sense of that phenomenon, given that there are alternative systems of set theory which *make sense* and which can model different things.

This objection would be spot-on, *if* I meant by 'set theory' some particular system of axiomatic set theory such as ZFC (or NF, etc.). But I do not. By 'set theory' I mean something much more basic and fundamental, which I call 'Chapter One' set theory. This is the view explained in the first chapter of textbooks (and the first lecture of courses) in every area of mathematics.[8] It is not so much a *theory* as a *way of thinking* with which we need to be inculcated if we are to understand any mathematical theory. Hence its placement at the beginning of books and lecture courses in all areas of mathematics, and hence also remarks such as the following:

> Mathematics habitually deals with "sets" made up of "elements" of various kinds, e.g., the set of faces of a polyhedron, the set of points on a line, the set of all positive integers, and so on. Because of their generality, it is hard to define these concepts in a way that does more than merely replace the word "set" by some equivalent term like "class," "family," "collection," etc. and the word "element" by some equivalent term like "member." We will adopt a "naive" point

> of view and regard the notions of a set and the elements of a set as primitive and well-understood. [Kolmogorov and Fomin, 1975, 1]
>
> We adopt, as most mathematicians do, the naive point of view regarding set theory. We shall assume that what is meant by a *set* of objects is intuitively clear, and we shall proceed on that basis without analyzing the concept further. [Munkres, 2000, 3]

These authors—and the others mentioned in note 9—then proceed to give some examples of sets, to introduce some terminology for sets and for notions such as membership, subset, intersection and union, and to work through some examples to give us facility with the new terminology and with how the various notions (subset, intersection etc.) interact with one another. Part of what is appealed to and reinforced in this process is precisely the idea that when we have some sets of objects, we automatically have all the facts about the identity and non-identity of their elements one with another. This is why—to take just one example—we can say that the intersection of sets $A$ and $B$ is the set of all objects which are in both $A$ and $B$. If there was a potential worry that it might not be fixed whether some object $a$, which we considered as a member of $A$, and some object $b$, which we considered as a member of $B$, are or are not the very same object, then the intersection would not be well-defined. Part of the view is, however, that the intersection is always well-defined (it might of course be empty)—and part of the necessary background for this is that whenever we have sets of objects, we automatically have a complete set of precise facts concerning the identity and non-identity of these objects one with another. (Of course we may not *know* these facts—we may not be able to tell, given a description of $a$ and a description of $b$, whether $a = b$—but the facts are all there nonetheless. If $A$ has been described to us as the set of solutions of a given equation, and $B$ as the set of even numbers, then we might not know what is in the intersection of $A$ and $B$—but a fundamental part of the basic view is that there *is* a particular set which is the intersection of $A$ and $B$, and that this set is fixed once the sets $A$ and $B$ are fixed.)

The very basic and fundamental view of sets laid out in Chapter One underlies *all* full-blown set theories: not just the standard alternatives (e.g., members of the ZF family, NBG, type theories), but also non-standard theories (e.g., Aczel's non-well-founded set theory $ZFC^{-}+AFA$), and indeed *anything* which we would recognise as a version of set theory. (In this connection, note that Chapter One set theory is *not* naive set theory. Naive set theory is a full-blown axiomatic set theory—or more accurately a family of such theories, whose common feature is that they include an unrestricted comprehension axiom. Chapter One set theory is a fundamental way of thinking which underlies all full-blown set theories, including naive set theory.) Thus the fact that vague identity is out of place in Chapter One set theory is a much deeper fact than the fact that the universal set is out of place in ZF. Vague identity is not merely not modellable in some specific system of set theory: it is not modellable in *anything* which we would recognise as a version of set theory. This is what underwrites my strong claim that we cannot make sense of vague identity at all.

**6.** At this point the friend of vague identity might see a way forward:

> OK, you have convinced me that those (such as Priest and van Inwagen) who attempt to model vague identity *within* anything like set theory as we know it are on the wrong track. To understand vague identity, we would need an entirely new framework for understanding things: we would require not just new models, but new modelling techniques. But perhaps such modelling techniques *can* be developed! What we should do is look for a radically new form of set theory—or something beyond set theory altogether—in which the idea of vague identity can be clearly understood. Do you have any reason to think we would fail in this project? Perhaps earlier efforts to make sense of vague identity have failed simply because they were aimed in the wrong direction: towards making standard set-theoretic models of vague relations, rather than towards developing radically non-standard set theories. Now that we see what needs to be done to make sense of vague identity, why think we cannot do what is required?

Of course no-one can conclusively rule out the possibility of future frameworks for understanding objects and collections in which sense can be made of the idea of vague identity. It would be absurd to think that we are now in a position to foresee every conceptual breakthrough that will ever be made. I can, however, say two things in response to the above line of thought.

First, the burden of proof lies firmly with the friends of vague identity to actually produce a framework in which we can clearly think about vague identity.

Second, while we cannot be absolutely sure that the friends of vague identity will fail in this quest, there is good reason to be sceptical about their prospects for success. For a start, we can certainly rule out the idea that past efforts to make sense of vague identity have failed simply because those involved were looking for the wrong thing (standard set-theoretic models of vague relations, rather than radically non-standard set theories). Krause et al. [2005, 232–4] write:

> standard set theories cannot deal with collections of 'genuine' indistinguishable objects. The standard way mathematicians consider indistinguishable things vary, but all of them can, in some way or another, be summed up by a technique used by H. Weyl to treat aggregates of individuals; in short, starting from a *set* $S$ with, say, $n$ elements, Weyl has assumed that there is an equivalence relation $R$ defined on $S$, and then he takes the equivalence classes $C_1, \ldots, C_k$ to play the role of collections of indistinguishable objects.... But all these 'solutions' are mathematical tricks, for the very characteristics of the elements of a set as *individuals* is always present, at least implicitly. So, this kind of devices cannot be considered as adequate answers to the *philosophical* problem of dealing with collections of

> indistinguishable objects.... from the philosophical point of view, it should be interesting to consider indistinguishability *right at the start*, as something which is very peculiar to the objects being supposed to exist, as it seems to be the case, in some situations, with quantum objects. In other words, if we take seriously the view that quantum objects shouldn't have individuality, that is, that they are to be taken as *non-individuals* in a sense, can we present a 'set theory' where indistinguishability is introduced right from the start?

It is clear that we and Krause et al. are on exactly the same page. Their discussion of Weyl makes the same basic point I made in my discussion of Priest, van Inwagen, and Trillas and Valverde (i.e., that overlaying a new relation on top of identity is not the same as making identity indeterminate), and they then state their goal of producing a radically new form of set theory (called 'quasi-set theory') which—unlike all standard set theories—does *not* regard objects as individuals, each of which is either determinately identical to or distinct from each other object.

So, do they reach their goal? No. Indeed they make no progress towards it. For they simply present quasi-set theory as an *axiomatic theory*—i.e., as a list of definitions and axioms involving a primitive relation of indistinguishability, symbolised as $\equiv$. This approach leads directly to the dilemma that Priest and van Inwagen faced. If we understand this axiomatic theory in the usual way—i.e., as picking out the class of (standard set-theoretic) models in which all the axioms are true—then we can only understand $\equiv$ as a new relation *in addition to* the precise identity relation which inheres in all standard models. We have then done nothing to *remove* precise identity in favour of the vague relation of indistinguishability—i.e., we have gone no way towards achieving Krause et al's stated goal of developing a framework which does *not* treat objects as individuals in the standard sense. Indeed we are in precisely the position of Weyl—whom Krause et al criticised precisely for being in this position. On the other hand, if we are not to interpret the axioms in the standard way, then we need to have a new way of understanding them explained to us—or else we do not understand them at all. Yet in spite of their opening statements (quoted above) Krause et al give no such explanation. From those opening statements, it sounded as though Krause et al would explain, from the ground up, a new framework for thinking about objects and collections in which we need not, and indeed must not, assume that whenever we have some objects, we have a complete and precise set of facts concerning the identity and non-identity of those objects one with another. Yet they do not do this: they simply present a set of axioms.

We have now seen that those who have set out to model vague identity have repeatedly ended up either with no clear model at all, or with a standard set-theoretic model of something *other than* vague identity—and this includes those who set out precisely to construct radically non-standard set theories, as opposed to standard set-theoretic models of vague relations. This strongly suggests (although it does not conclusively establish) that our current way of thinking about identity is not something like a *convention*, to which there exist alternatives, but is a fundamental part of the background framework which we *use* whenever we try to think clearly *about* anything whatsoever. To elaborate on this distinction: Except in places such as the exits of the tunnel between France (where they drive on the right) and England (where they drive on the left), we do not have signs on our roads telling us to drive on the left or right. This does not mean that there is something *deeply* special about the side of the road we drive on—that we *could not* all decide to drive on the other side. It just means that the convention is so firmly entrenched that we do not need to be constantly reminded of it. Contrast this with a different sort of case. It is plausible to say that reasoning inductively from a sample is not a convention we have (tacitly) adopted: it goes much deeper than that. We have no serious alternative to induction: no alternative that we can think *with*, as opposed to think *about*. That is, even when we think about counter-inductivists, we cannot help thinking inductively ourselves. What I want to suggest is that the situation with regards to our current thinking about identity is like the induction case, not the driving case. It is not that our current practice of making free appeal to a classical identity relation whenever we have a set of objects is a convention which is so firmly entrenched that we do not need to state it explicitly. Rather we *have no alternative*. This way of thinking is the only way we have—and will ever have—of thinking clearly about objects and collections of objects. We can think *about* other ways of thinking about identity, but we cannot think *with* them—we cannot accept that they are genuinely alternative ways of thinking about *identity*—for, as we have seen, the conception of sets of objects as pre-equipped with the classical identity relation is always operative in the background.

**7.** I said at the outset that my argument against vague identity is more fundamental than existing arguments. It is time to substantiate this claim. I shall do so not by considering every existing argument against vague identity individually—that would take far too long—but by considering prominent examples of the two broad classes of argument to be found in the literature: formal arguments and philosophical arguments.

The most prominent formal argument against vague identity is that of Evans [1978], which runs as follows. Suppose (for purposes of *reductio*) that it is indeterminate whether $a = b$. Let us express indeterminacy by the symbol $\nabla$. Then we have:

$$\nabla a = b \tag{1}$$

So $b$ has the property of being indeterminately identical to $a$:

$$\lambda x[\nabla a = x]b \tag{2}$$

But it is not indeterminate whether $a = a$:

$$\sim \nabla a = a \tag{3}$$

So $a$ does not have the property of being indeterminately identical to $a$:

$$\sim \lambda x[\nabla a = x]a \tag{4}$$

So there is a property which $b$ has and $a$ does not; so $a$ is not identical to $b$:

$$a \neq b \tag{5}$$

According to Evans, this contradicts the initial supposition that $\nabla a = b$.

How are we to assess this argument? We need to know two things: whether it is valid; and whether the conclusion really does contradict the initial supposition. In order to know these things we need a semantics or model theory for the propositions in the argument. Now, of course, the friends of vague identity will offer model theories in which the argument is not valid,[9] or in which $\nabla a = b$ and $a \neq b$ can both be true. And now what are we to say? The proponent of Evans's argument will need to do two things: provide her own model theory, which validates the argument and renders it a true *reductio*; *and* convince us that it is better than alternative model theories which do not validate the argument, or do not render it a true *reductio*. I can see no way of doing this other than pointing out to the friends of vague identity either that their models are not explained properly (if they are not models of the standard set-theoretic sort), or that what they are calling vague identity in their models is not really identity. That is, there is no way of getting any solid result out of Evans's argument without appealing to the argument against vague identity in this paper.[10]

Turning from formal to philosophical arguments, Williamson [2002][11] has recently advanced an argument against vague identity that has a strong affinity with part of my argument in this paper.[12] Like me, Williamson argues that the friends of vague identity make the mistake of interpreting '=' as something other than identity, when they present their model theories. Unlike me, Williamson focusses on the relationship between classical metalanguages and vague object languages. He argues that if the metalanguage in which the friend of vague identity presents her model theory is classical, then either she fails to provide a model of *identity*, or she models identity as *precise* rather than vague:

> on classical metalogical assumptions, either the denotations of $a$ and $b$ in a model $M$ are identical or they are not. If they are identical and $M$ is faithful to

> the intended reading of $=$ then $a = b$ is true in $M$; if the denotations are not identical and $M$ is faithful then $a = b$ is false in $M$. Either way, $a = b$ is bivalent in $M$. To avoid this result, the non-classical models treat $=$ unfaithfully [i.e., as meaning something other than genuine identity]. [Williamson, 2002, 283]

Williamson then continues:

> Those remarks are not intended to show by themselves that there can be no faithful model of indeterminate identity, for they do not show that the metalogic must be classical. Someone might... by non-classical reasoning in the metalanguage still avoid conceding that $a = b$ is either true or false in the model... But that move would be as controversial as their non-classical metalogic. The coherence of vague identity would not have been established. [Williamson, 2002, 284]

Both parts of the argument are skeletal—and what is needed to flesh them out are precisely the arguments of this paper. In the first part of the argument, Williamson makes free appeal to a classical identity relation in the model under consideration, without discussion of *why* we may always assume the existence of such a relation. He thus has no answer to Priest's complaint that "[t]he objector does not even have a right to *assume* that the domain $D$ is furnished with a relation of the kind required for classical identity" [Priest, 1998, 337]. Enter this paper, with its detailed explanation of why we are entitled to assume the existence of a classical identity relation on any set of objects. In the second part of the argument, Williamson says that using a non-classical metalanguage would be controversial, but gives no justification for his negative attitude towards non-classical metalogic. Without such a justification, his argument has no force against the friends of vague identity: it merely shows them the path they must pursue, without giving any reason to think the path impassable. Again, enter this paper: given what has been said above, we can see precisely what would be wrong with presenting a non-classical model theory for vague identity, and then saying that the metalanguage *in which* the model theory was presented is non-classical in the same way as the object language *for which* the model theory was given.[13] The problem is that we only understand the model theory for vague identity in the first place if we take it to be a piece of standard mathematics—a set-theoretic construction of the standard sort. So if the friend of vague identity turns around at the end of her presentation of this model theory and tells us that the language in which she made her presentation was governed by the very semantics she just presented, then we have to conclude that we did not understand the presentation at all. We are back at square one: we thought she was presenting a piece of standard mathematics, and we know how to understand that sort of thing; but if she was not, then unless some other way to understand the presentation is explained to us from the ground up, we do not understand it at all. As discussed above, no friend of vague identity ever has explained such an alternative way of understanding things, and I have given reasons for thinking that they never will.[14]

## Notes

[1] Thanks to John Cusbert for helpful comments on and discussions of an earlier draft—in particular for pressing me in relation to the imagined response to my argument discussed in §6; to Amitavo Islam for observations about how mathematicians make sense of new concepts, which were an important influence on the view of §2; to Graham Priest for helpful correspondence about his paper, and for the paper itself—it was close reflection on his view that led me to the central argument of the present paper; and to the anonymous referee for several helpful comments.

[2] For more examples, see Parsons [2000].

[3] In the first case I presented, although the cloud-mass changes over time, the questionable identity is between clouds L and R at one time (a time in the middle of the process of change). The second and third cases, on the other hand, involve questionable identity between $a$ at one time and $b$ at another time. My discussion in this paper is concerned with identity *simpliciter*—whether between $a$ and $b$ at one time, or between $a$ at one time and $b$ at another.

[4] Real examples from the literature will be discussed below.

[5] Priest has a footnote at the end of the second sentence in the passage quoted above: "In the mathematical theory of metric spaces, metrics are usually taken to satisfy the extra condition $d(x, y) = 0 \Rightarrow x = y \ldots$ Whilst this condition could be added here also, it plays no significant part in what follows" [Priest, 1998, 340, n.5]. In fact this condition could not be added here without immediately ringing alarm bells: for it makes reference (on its right hand side) to the real identity relation on the domain—and if this relation is there, then it will, as just discussed, trump Priest's proposed interpretation of '=' as referring not to this real identity relation, but to the metric structure. Of course this is not to say that leaving this condition out avoids the trouble (it only avoids the alarm bells). My whole point is that as soon as we have the domain, we have the identity relation thereon. No invocations are required to put it there, and no invocations (or lack thereof) will remove it.

[6] See also the first and second paragraphs on their p.231, and their p.242.

[7] Indeed we can do so in many different ways: any theory of vagueness which gives a model theory for vague language of the standard set-theoretic sort (e.g., supervaluational models, fuzzy models, three-valued models, ...) provides such a way. (To ward off a potential confusion: I do not mean that fuzzy (supervaluational, etc.) models are classical models; I mean that they are constructed using standard set-theoretic tools—that fuzzy (supervaluational, etc.) model theory is part of standard mathematics.)

[8] Picking a few books from my shelves, I find for example Mac Lane and Birkhoff [1999, ch. 1], Kolmogorov and Fomin [1975, ch. 1], Hoffman and Kunze [1971, Appendix] (although their material on sets occurs in an appendix rather than in the first chapter, Hoffman and Kunze note that it "is more in the nature of an introduction for the book than an appendix" [386]), Chung [1979, ch. 1] and Munkres [2000, ch. 1].

[9] For example Priest [1998] and van Inwagen [1988].

[10] Similar comments apply to the arguments of, amongst others, Salmon [1979, 243–6] and Pelletier [1989].

[11] See also Williamson [2003, §8].

[12] I discovered Williamson's views after formulating my own.

[13] The problem would be just as bad if we said that the metalanguage and the object language were non-classical in *different* ways.

[14] At the end of his illuminating paper on Evans's argument, Heck [1998, 292–3] clearly envisages the *sort* of thing that the friend of vague identity needs at this point—namely a

*new type* of semantic theory. Heck does not, however, attempt to provide such a theory in his paper.

## References

Kai Lai Chung. *Elementary Probability Theory with Stochastic Processes*. Springer-Verlag, New York, third edition, 1979.

Gareth Evans. "Can there be vague objects?" *Analysis*, 38:208, 1978.

Richard G. Heck. "That there might be vague objects (so far as concerns logic)." *Monist*, 81:274–96, 1998.

Kenneth Hoffman and Ray Kunze. *Linear Algebra*. Prentice Hall, Upper Saddle River NJ, second edition, 1971.

A. N. Kolmogorov and S.V. Fomin. *Introductory Real Analysis*. Dover, New York, revised English edition, 1975. Translated and edited by Richard A. Silverman.

Décio Krause, Adonai S. Sant'Anna, and Aurélio Sartorelli. "On the concept of identity in Zermelo-Fraenkel-like axioms and its relationship with quantum statistics." *Logique et Analyse*, 189–192:231–60, 2005.

Saunders Mac Lane and Garrett Birkhoff. *Algebra*. AMS Chelsea, Providence RI, third edition, 1999.

James R. Munkres. *Topology*. Prentice Hall, Upper Saddle River NJ, second edition, 2000.

Terence Parsons. *Indeterminate Identity: Metaphysics and Semantics*. Clarendon Press, Oxford, 2000.

Francis Jeffry Pelletier. "Another argument against vague objects." *Journal of Philosophy*, 86:481–92, 1989.

Graham Priest. "Fuzzy identity and local validity." *Monist*, 81:331–42, 1998.

Nathan Salmon. *Reference and Essence*. Princeton University Press, 1979.

E. Trillas and L. Valverde. "An inquiry into indistinguishability operators." In Heinz J. Skala, S. Termini, and E. Trillas, editors, *Aspects of Vagueness*, pages 231–56. D. Reidel Publishing Company, Dordrecht, 1984.

Peter van Inwagen. "How to reason about vague objects." *Philosophical Topics*, 16:255–84, 1988.

Timothy Williamson. "Vagueness, identity and Leibniz's law." In Andrea Bottani, Massimiliano Carrara, and Pierdanielle Giaretta, editors, *Individuals, Essence and Identity: Themes of Analytic Metaphysics*. Kluwer, Dordrecht, 2002.

Timothy Williamson. "Vagueness in reality." In Michael J. Loux and Dean W. Zimmerman, editors, *The Oxford Handbook of Metaphysics*. Oxford University Press, 2003.