ABSTRACT. In this paper it is argued that The Principle of the Identity of Indiscernibles can be justified as a concrete application of Ockham's Razor, the maxim which enjoins us not to multiply entities beyond necessity. First, a statement of the Principle is presented, according to which the Principle, while interesting enough, is not logically necessary. It is then argued that the assumption of the falsity of the Principle prescribes an epistemological situation where it seems to be impossible to find grounds for thinking that the Principle is indeed false. Hence it is to be accepted as an epistemological necessity of sorts, one that recommended by the desire not to multiply entities beyond need.

In this paper I argue that we can justify the acceptance of The Principle of the Identity of Indiscernibles (PII) as a concrete application of Ockham's Razor, the maxim which enjoins us not to multiply entities beyond necessity.

In § 1 I distinguish between (PII) and certain other principles, or rather, interpretations which often pass for (PII), but which, I think are better called something else. (PII), as I view it here, is interesting enough, but it does not follow from logic, or from facts regarding the occupation of places at times.

In § 2 I offer an argument to show that the assumption of the falsity of (PII) prescribes an epistemological situation where it seems to be impossible to find grounds for thinking that an object a and an object b are distinct yet indiscernible. That is, I aim to show that we shall never have to think that there are distinct yet indiscernible objects. Once this is established, we can proceed to apply Ockham's Razor, by declaring the identity of the "distinct" indiscernibles.
§ 1, PRINCIPLES OF THE IDENTITY OF X-ICALS

Socrates is identical with Socrates. But does he have a property of being identical with Socrates? And, in general, does a thing \( x \) have a property of being identical with \( x \)? These are controversial questions whose answer depends on how the term 'property' is understood. Having no wish to deal with these questions, let me nevertheless assume, for the sake of distinguishing different ways of taking (PII), that there are such properties as being identical with \( x \), where \( x \) is any particular object.

Without any question, anything \( y \) that has the property of being identical with \( x \) just in case \( x \) has it, is none other than \( x \). To express this truism in symbols we write:

\[
(1) \quad (x)(y) (x=y \supset x=y)
\]

The antecedent of this conditional says really little more than just \( x = y \), or, in accordance with our present assumption, that \( y \) has the property of being identical with \( x \). This being so, we shall honor (1) with the imposing name 'The Principle of the Identity of Identicals'. For it simply says that no two things are identical, or, in somewhat more impressive terms, that no two things share an identity property, which is what we take the sentential function \( x = (\ldots) \)' to designate.

Along similar lines, let us suppose that each thing \( x \) has a property of occupying space-time region \( p-t \), or being at \( p-t \), for short. Assuming, furthermore, that no two things can occupy the same space-time region, we lay down the principle that

\[
(2) \quad (x)(y) (x \text{ is at } p-t \supset y \text{ is at } p-t. \supset x=y)
\]

To this principle we give another imposing name: we call it 'The Principle of the Identity of the Spatia-Temporally Coincident'. It simply says that no two things occupy the same space-time region, or, in somewhat different words, that no two things share a spatio-temporal property, which is what we take the sentential function '(\ldots) is at \( p-t \)' to designate.

In general, if each thing \( x \) has a property \( X \) which is such that nothing but \( x \) can have, then we are in position to speak of a Principle of the Identity of X-icals, a principle whose general form is:

\[
(3) \quad (x)(y) [X(x) \ldots X(y). \supset x=y]
\]

A further illustration of this form may be found in Russell's and Goodman's notion of spatio-temporal "qualia" or visual space locations. (See Russell [9], 298-9, and Goodman [5] 196 and 201.)

Now over and above Principles of the Identity of X-icals stands the familiar Principle of the Identity of Indiscernibles. This principle may be represented thus:

\[
(4) \quad (x)(y) [(f) (fx = fy) \supset x=y]
\]
This principle does not single out any specific properties for special attention. It simply says that no two things share all their properties, without suggesting that there are any specific properties which cannot be shared.

Still, depending on how broad one's conception of properties is, it is possible to view (4) so that it is no more than an uninteresting generalization of some pet Principle of the Identity of X-icals. Thus if someone believes that no two things can share an identity property then, a fortiori, he will believe that no two things can share all their properties. Whatever interest (4) has then, become secondary, deriving from whatever interest the associated Principle of the Identity of X-icals has.

In what follows I do not intend to criticize the views which base (PH) on some underlying Principle of the Identity of X-icals. This is not because I am sympathetic to such views; in fact, I do not think that it is right to refer to (PH) by this name when it is understood to be trivially derivable from principles that do not even mention the word 'indiscernibility'. Rather, it is because it is very hard to say what is to count as a property.

Thus what I am proposing is to take (PH) so that it is a basic principle, or, at any rate, one that is not derivable from less general principles. And this requires that we exclude from consideration all the properties whose having serves of necessity to make an object unique.

This leaves us with properties each of which can be had by a multiplicity of objects. And while it is possible to make finer distinctions within this class of relevant properties (obviously, there are many differences between being white, being a man, and being a son of Adam), still this will not be necessary for purposes of the argument which we will be presenting in support of (PII). Suffice it to say that every property which we consider is one that can be had by an indefinite number of objects.

We can now ask: What happens to (PII) when it is taken the way we take it here? The first thing that happens, it is commonly acknowledged, is that (PII) becomes logically contingent. Even Leibniz, whom we have just quoted as saying that two indiscernibles would be two, continues by saying that such a state of affairs "would be against the grand principle of Reason", by which, I take it, he means The Principle of Sufficient
Reason. And in the many recent discussions of (PII) there has been no dearth of suggestions to the effect that we can redescribe the universe, identifying the distinct indiscernibles, rather than give up belief in (PII). A brief glance at some of these suggestions will enable me to state what I find to be lacking in them, as well as pave the way for introducing my own defense of (PII).

M. Dummett makes the point in favor of (PII) briefly and clearly when he says:

If it is really the case that we can find no predicate not itself involving reference to either an object \( a \) or an object \( b \), which is true of \( a \) but not of \( b \), then nothing can possibly form an obstacle to our regarding \( a \) as identical with \( b \). If, for example, we suppose a completely symmetrical universe consisting of two spheres and an observer (himself always symmetrical about the plane of points equidistant from the two centres), nothing can stand in the way of a redeescription of the universe as consisting of a single sphere and an observer (that is, so to speak, half an observer). ([4], 543-4)

Dummett's suggestion here is relevant to the case of what might be called "spatial duplication"—the duplication of things in space. A. Grünbaum similarly deals with what might be called "temporal duplication"—the duplication of events in time. He asks us to imagine a universe consisting of a lonely platform, on which a particle moves in a circular path. Shall we say that there is a type of event, e.g., the particle's being at the place marked A, which keeps recurring eternally? No. Grünbaum's suggestion is that

Instead of appearing periodically at the place A at different instants in open time, the particle would be "returning"—in a highly Pickwickian sense of that term—to the self-same event at the same instant in closed time. This conclusion rests on Leibniz's thesis that if two states of the world have precisely the same attributes, then we are not confronted by distinct states, but merely by two different names for the same state at one time. ([6], 197)

Like Grünbaum, B. van Fraassen ([11], 65) and W.H. Newton-Smith ([8], 74-6) think that it is possible for one to insist upon (PII) by describing, or, if you wish, redescribing, the world's time as being closed.

The motive behind such suggestions, it seems to me, is none other than the commendable motive of ontological economy, based on the tacit assumption that distinct indiscernibles are not needed for purposes of explaining phenomena, or for achieving an adequate theory of the world. For if it were to be thought that there could be situations where we somehow had to acknowledge both distinctness and indiscernibility, then it would be far from reasonable to suggest that we redescribe, identifying the indiscernibles. Therefore it must be thought that we can do without distinct indiscernibles. Ian Hacking makes a good point when he says, on behalf of
the Leibnizians, that "Richness ... comes from a variety of phenomena, not from iteration of the same thing". ([7], 253)

This being so, we must have an argument that makes it reasonable to think that there will never arise situations where we might have to acknowledge both distinctness and indiscernibility.

It is not obvious how such an argument would proceed. Presumably, it would have to do with the fact that the objects said to be distinct are also said to be indiscernible. But are we not then assuming, as Max Black puts it ([3], 214) that "... in order to verify that there are two things of a certain kind, it must be possible to show that one has a property not possessed by the other"? If so, then we must be prepared to answer an objection which Black expresses in this manner:

We can verify that there are two [objects of a certain kind], that is to say, a certain property of the whole configuration, even though there is no way of detecting any character that uniquely characterizes any element of the configuration. ([3], 215)

The objection is certainly not an idle one. Any suggestion to redescribe the universe will be premature unless we first give reasons for thinking that the situation which Black abstractly envisions will never arise.

Hacking is one philosopher who comes to grips with this problem if only in a partial way. In [7] he discusses how a Leibnizian supporter of (PII) and a Kantian opponent might argue over Kant's example of two indiscernible drops of water. Various things happen in the universe of the solitary water drop(s), the Kantian describing things one way, his Leibnizian opponent describing them another. Hacking exercises a lot of ingenuity on behalf of the Leibnizian, and he seems to think the latter can handle anything that the former can come up with.

But what we really need is a general argument for thinking that the Leibnizian can always do that. Might it not be possible for a Kantian to come up with an example that would finally reduce the Leibnizian to silence? How can we be sure that future experience will not confront us with a situation where we have to acknowledge both distinctness and indiscernibility?

In what follows I offer an argument which, if sound, will show that we need not fear that such a situation will ever arise.

§2 INDISCERNIBILITY AND JUDGMENTS OF DISTINCTNESS

The present argument in favor of (PII) may be briefly stated by saying that the assumption that an object \( a \) and an object \( b \) are indiscernible robs us of all epistemological grounds for thinking that \( a \) and \( b \) are numerically distinct. Nothing will make it necessary for us to believe that \( a \) and \( b \) are distinct, and, hence, it will always be possible for us to apply Ockham's Razor to declare their identity.
The argument is presented with the help of a historical detective fiction of sorts. This will enable us to talk about (PH) in the "formal mode", which will be rather convenient. Later, it should not be difficult for us to translate our conclusions into the "material mode", if we want to.

Let us imagine that historians one day discover an ancient manuscript which purports to tell the life-story of a man called 'Opp' and a man called 'Oss'. We are to read the manuscript and determine, if we can, whether 'Opp' and 'Oss' are names of two distinct men, or two names of one and the same man.

For our convenience, the manuscript is divided into three parts. Part I describes Opp's life: it contains all the predicates that were ever true of him, with the exception of those that make reference to Oss. Part II does the same for Oss: in it we find all the predicates that were ever true of him, with the exception of those that make reference to Opp. Finally, Part III brings Opp and Oss together: in it we find the relational predicates which apply to Oss in relation to Opp, and to Opp in relation to Oss.

But this is not all. To make our detective work a little interesting, we shall imagine that whoever wrote the manuscript had taken care to omit from Part III all the relational predicates which may be thought to beg the question for or against the identity of Opp and Oss. For example, we do not want Part III to contain statements such as "Oss = Opp", or "Oss ≠ Opp". Nor would it be fair to have a statement such as "The body of Oss = the body of Opp", assuming, as we normally do, that two men do not have one body. Nor, to give one last example, would it be fair to have a statement such as "Opp lived 10 miles away from Oss", for this means, in a question-begging way, that the place of Opp is distinct from that of Oss. In general, all objects which would be one (or two) if and only if Opp and Oss were one (or two) must be excluded. Neither the identity nor the distinctness of such objects may be asserted.

With this restriction on the material to be found in Part III we can proceed to ascertain whether Opp and Oss are distinct. First, we read Part I, which gives an Oss-free description of Opp's life. Then we read Part II, which gives an Opp-free description of Oss's life. And here we find, to our great surprise, that everything which Opp did, or was, applies to Oss, and vice versa. In fact (we find), Parts I and II are exactly the same, except that whereas one has 'Opp', the other has 'Oss'.

This, very naturally, suggests to us that Opp and Oss are really one man who somehow managed to have his story told twice. But, of course, opponents of (PH) will tell us to go on to read Part III. For there we might come across a relation whose holding between Opp and Oss could give us grounds for thinking that Opp and Oss are distinct. With Black, one might say that a property of the "whole configuration" (Opp in relation to Oss) might convince us of numerical distinctness, even if we cannot find a property that applies to one but not the other.

Is it possible to find a relation of this sort in Part III? I do not think so. Imagine that we begin to read this part. Whenever we come across the name 'Opp' we
delete it and write 'Oss'. Thus according to our rewriting project, all relations which
hold between Opp and Oss are to hold between Oss and himself (or Opp and himself,
for it does not really matter which name we delete). Can we expect to be able to re­
write Part III in its entirety in this fashion? Might we not come across a relation R
which is such that a thing cannot have to itself? It is all right for our rewriting project
to yield such statements as "Oss sees Oss", and "Oss likes Oss", for Oss can see and
like himself. But can we expect this to be the case with all the relations that can hold
between Opp and Oss?

This is the difficult question upon whose answer our whole argument de­
pends. The answer is Yes. But before we proceed to make a case for this answer, we
need to state clearly and emphatically two conditions which must be met by any rela­
tion R that can be said to prove us wrong.

Firstly, what we need is an irreflexive relation R, one that cannot hold be­
tween something and itself. If R could, then it would be possible for us to claim, self­
consistently, that R holds between Oss and himself (or Opp and himself.) And this, of
course, would defeat the purpose of establishing that Opp and Oss are distinct.

Secondly, we must require R to be symmetric, that is, if anything x has R to
anything y, then y has R to x. If this were not the case for the relation we want, then
Opp and Oss would cease to be indiscernible. To prove this, assume: R (Oss,Opp).
but not R(Opp,Oss). Here then is a property which Oss, but not Opp, has. Supposing
that F is the conjunction of all the shareable properties which Opp and Oss have in
common, Oss can be distinguished by being R to someone who is F. Opp does not
have this property. He is not R to someone who is F. After all, who could this be?
Not Oss, for this conflicts with the assumption of the present argument. And it can­
not be Opp himself, for, as we have just said, R cannot hold between something and
itself. (It won't help to say: Someone else then. We do not need to assume that there
is anything else for Opp and Oss to be related to. As far as establishing the truth or
falsity of (P11) is concerned, Opp and Oss could be the only thing(s) in the universe.)

We may now state, in summary, that we are looking for a relation R that
would satisfy these two conditions:

(5) (x)¬R(xx) \hspace{1cm} \text{(Irreflexivity)}

(6) (x) (y) (R(xy) ∨ R(yx)) \hspace{1cm} \text{(Symmetry)}

Now suppose we do find an R of the sort we are looking for. Whether R
holds from Oss to Opp (that is, R(Oss,Opp)), or from Opp to Oss (that is,
R(Opp,Oss)), in either case, we have, by obvious logical inferences from (5) and (6):

(7) \hspace{1cm} (a) R(Oss,Opp)
(b) ¬R(Oss,Oss)
(c) R(Opp,Oss)
(d) ¬R(Opp,Opp)
On the face of it, (7) does not look problematic. Nevertheless, I think it can be shown to be an epistemological nightmare when taken in conjunction with the fact that Opp and Oss share all the properties which are mentioned in Parts I and II of our manuscript.

To show this, consider what it would be like for us to have evidence that \( R(\text{Oss}, \text{Opp}) \). (An exactly parallel argument applies when we suppose that we have evidence that \( R(\text{Opp}, \text{Oss}) \).) When I know that \( \text{Oss} \) is \( R \) to \( \text{Opp} \), I know \textit{ipso facto} that \( \text{Oss} \) is not \( R \) to \( \text{Oss} \). As we have said, \( R \) is the sort of relation that cannot hold between something and itself. But how am I to know that \( \text{Oss} \) is not \( R \) to \( \text{Oss} \), unless I knew beforehand that \( \text{Opp} \) and \( \text{Oss} \) are distinct? Suppose I did not know that for sure. Would I then know that \( \text{being } R \text{ to } \text{Opp} \) applies to something (anything) that \( \text{being } R \text{ to } \text{Oss} \) does not apply to? Surely not. To know that \( \text{Oss} \), in being \( R \) to \( \text{Opp} \), is not as a matter of fact \( R \) to himself, I must know beforehand that he is not \( \text{Opp} \). Otherwise, \( \text{Oss} \), in having the one condition of \( \text{being } R \text{ to } \text{Opp} \), has the other (\( \text{being } R \) to \( \text{Oss} \), himself, that is.) And this obviously means that the irreflexive character of \( R \) has been compromised. Not only is \( \text{Oss} \) \( R \) to \( \text{Opp} \), but for all we know, he is \( R \) to himself.

Of course, if \( \text{Oss} \) and \( \text{Opp} \) had different properties on their own (these would have appeared in Parts I and II of the manuscript), then we would have had our "prior" knowledge of distinctness, and we would not be now puzzling over the question of whether \( \text{Oss} \), in being \( R \) to \( \text{Opp} \), is actually \( R \) to himself. But, unfortunately, \( \text{Opp} \) and \( \text{Oss} \) do not differ on their own. Consequently, (7) becomes unknowable.

If the above argument is sound, then we can be certain that we shall never find in Part III an \( \text{Opp-Oss} \) relation that shows, beyond doubt, that \( \text{Opp} \) is not \( \text{Oss} \). Therefore we should be able to rewrite Part III, replacing '\( \text{Opp} \)' by '\( \text{Oss} \)' throughout. Of course, in so doing, we apply Ockham's Razor.

Similar reasoning should, I believe, apply to Kant's solitary water drops, to Strawson's lonely chess-board (See [10], 120) and to Black's more famous spheres, Castor and Pollux.

Taking the latter as our example, imagine that we keep a Castor File which contains all that we know about Castor, except for items of knowledge that make reference to Pollux. Also let a Pollux File contain all that we know about Pollux, except for items of knowledge that make reference to Castor.

If the two files fail to distinguish between Castor and Pollux, how can we be sure that these are really two? Black suggests, in general terms, that Castor and Pollux might together "do" something, a cumulative effect of sorts, which will convince us that they are really two.

Let us, therefore, open a Castor-Pollux File, ready for the reception of any items of knowledge that involve reference to both. What can possibly go into this file, which we cannot reinterpret as involving only Castor, or only Pollux? Of course, we exclude question-begging items such as 'Castor is other than Pollux', and 'Castor is a hundred miles away from Pollux'. What relation \( R \) of any description or complexity,
can be known to hold between Castor and Pollux, but which we cannot self-consistently claim to hold between Castor and itself (or Pollux and itself)? None, if our argument using Opp and Oss is sound. In fact, we can rewrite the relevant parts of the preceding discussion replacing 'Opp' by 'Castor', and 'Oss' by 'Pollux'. The conclusion would still be the same: nothing can prevent us from declaring Castor and Pollux to be identical.

But we must be careful not to read too much into this agreeable result. Specifically, we must not think that we have shown that it is always possible to identify distinct indiscernibles in such a way that the resulting view of the facts is in every way a satisfactory view. Thus it may be that upon identifying the distinct indiscernibles we find that we have made space, or time, closed. Or it may be that the upholding of (PH) will require giving up, or making changes in, a theory that we are in many ways happy with. Thus the acceptance of (PH) may come at a price.

Nevertheless, I do not think that this, by itself, can be allowed to count as a good argument against accepting (PII). The reason is simply that the rejection of (PII) also comes at a price. To illustrate this by means of an easy-to-follow-example, consider what it must be like to live in a universe where you have an indiscernible double. You cannot talk with your double, for he will have to say the same things as yourself. You cannot play games, or do anything interesting with your double. In fact, he would be as good company to you as your mirror image. Surely this is not an agreeable consequence, considering that your double is every bit as real as you are.

Real entities, one thinks, must be able to act on one another. But a world with indiscernible entities is a world where the possibilities of interaction are severely limited by the fact that no asymmetric action can take place between indiscernible doubles. Thus one cannot (even in principle) defeat one's indiscernible double at a game of chess. Nor can a snake devour its indiscernible double.

One begins to wonder what justification there can be for thinking that indiscernible doubles inhabit the same world. For the unity of the world does not derive only, or even primarily, from the fact that everything in the world is somewhere in space-time, but also from the fact that everything can interact with everything else. A world with indiscernible doubles is a deeply divided world. And surely, this is not a welcome consequence.

Hence it is no argument to say that the adherence to (PII) can lead to undesirable consequences. For so can its denial. In addition, when we consider the point of ontological economy, the latter becomes a case of too much of a bad thing.
REFERENCES


