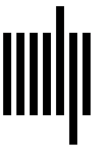
Time and Identity

Edited by: Joseph Keim Campbell, Michael O'Rourke, Harry S. Silverstein

Citation:

Time and Identity
Edited by: Joseph Keim Campbell, Michael O'Rourke, Harry S. Silverstein
DOI: 10.7551/mitpress/9780262014090.001.0001
ISBN (electronic): 9780262265799
Publisher: The MIT Press
Published: 2010



The MIT Press

10 Me, Again

Jenann Ismael

I traveled to the conference from which this volume is drawn with John Pollock, a dear and now departed friend and colleague. I dedicate this paper to him with love and respect. He is missed.

Introduction

Thought about the self raises some very special problems. Some of these concern indexical reference quite generally; but there is one, having to do with identity over time, that seems to be unique to the self. I'll be using a historical exchange between Anscombe and Descartes to raise the problem and proposing a resolution that casts light both on why self-directed thought presents a unique epistemic predicament and where Descartes's *cogito* argument—still one of the most compelling and resilient arguments in philosophy—goes wrong.

What Anscombe Said to Descartes

Descartes begins his discussion in the *Meditations* with the question "what am I?" and concludes, famously, that he is a nonmaterial substance. His reasoning turns on the thesis that nothing can be true of his nature that is not made known to him in the act of thinking.[1]

Schematically, the argument runs thus:

(1) I $=_{def}$ that thing whose existence cannot be doubted (i.e., that thing whose existence is made known in the act of doubting that it exists).

(2) I can't be identical to anything whose existence can be doubted.

(3) I can doubt the existence of anything for which objective criteria of identity can be provided (give me a description and I can coherently doubt that it is satisfied).

---

(4) Hence, I can't be identified with anything for which objective criteria of identity can be given. (Nagel 1989, 35)[2]

So construed, the argument is subject to an objection—often associated with Lichtenberg, but repeated by Russell, and given its most articulate expression by Anscombe—that is, on first encounter, devastating (Anscombe 1990).[3] The charge is that Descartes's argument fails because he pulls a bait and switch, overtly declaring the object of his inquiry is nothing other than that whose existence is made known in the act of trying to deny that it exists, but then tacitly appealing to a richer notion in allowing that it can be reidentified in different thoughts. In Anscombe's words:

> People have sometimes queried how Descartes could conclude to his *RES cogitans*. But this is to forget that Descartes declares its essence to be nothing but thinking. . . . His position has, however, the intolerable difficulty of requiring an identification of the same referent in different I-thoughts. (Anscombe 1990)

It is not the proclaimed guaranteed existence of the self that is problematic, for that does indeed follow if the 'I' in an 'I'-thought is purely reflexive, that is, if it simply refers to itself. It is the combination of guaranteed and *continuing* existence that is objectionable. For, if the thinker outlives the thought—if the 'I' in one thought is even *potentially* intersubstitutable with the 'I' in others—its existence can't be revealed in the mere production of an 'I'-thought. All whose existence is so guaranteed are the individual thoughts themselves. According to the argument, Descartes is not entitled to assume the existence of an underlying continuant to supply them a common referent. On the one hand, the concept needs to be kept bare to guarantee existence. On the other hand, the bare concept doesn't seem to be enough to support reidentification. If I cannot be something whose existence is not made known in the production of an 'I'-thought, something whose existence can be coherently doubted while I am thinking of it, then I cannot be something that even potentially recurs. The unadorned reflexive 'I' of the individual thought cannot pull other 'I' occurrences under its referential scope.

'I' and Other Indexicals

The difficulty can be brought out with a comparison of 'I' with names and other indexicals. In the case of names, ambiguity aside, criteria for identifi-

cation and individuation of words substitute for criteria of identification and individuation of objects.[4] The user of what Frege used to call a "logically perfect language" with a name for every object and no indexicals could be assured that the same word used on two occasions or in two expressions referred to the same object and could be intersubstituted *salva veritate*. She would never need to look beyond language to be assured of the validity of inferences or substitutions. This purity is spoiled when indexicals are added to a language. The user of an indexical uses the same word to refer to separate objects on different occasions and has to look to the world to disambiguate. Reidentification and intersubstitution requires knowing what type of object the expression picks out, and one has to apply criteria of identity associated with the type to determine when reidentification has been achieved or when intersubstitution is licensed. Spatial and temporal indexicals, for example, require application of criteria of identity for places and times, respectively. To use an expression like 'here' you have to know that 'here'-occurrences refer to places, and to make valid substitutions you have to be able to recognize that a place identified on one occasion is the same place as that identified on another.[5] Cases of demonstrative reference that we understand best have the form 'this *x*' and combine a reflexive component with a sortal. With indexicals, the sortal is part of the semantic meaning of the term. With demonstratives, it is either supplied verbally (as in 'this car' or 'that cup of coffee') or made clear enough in some other way by the conversational context (as when I say "that" pointing at a bookshelf and we both understand without saying that it's a particular book that I mean to be indicating). Reference, these cases, involves a division of labor. The reflexive component centers reference on a particular time, place, or object, and the sortal provides criteria for reidentification and intersubstitution.[6] Even when we don't know exactly what we're pointing at, we can say something like "that black hulk moving across the field," letting criteria for individuation of black hulks decide whether I've got my eye on the *same* thing that you do. If we adopted a Kaplanian semantics according to which demonstratives are functions, we would say that the sortal proscribes the range and the reflexive component picks out the value (Kaplan 1989a,b).

To fit 'I' into this mold, we would have to say that occurrences of 'I' refer to the selves that produce them and intersubstitution of 'I' in temporally separated thoughts involves application of criteria of identity for selves. We would have to say that each of us applies criteria of identity for selves in judging that some past or future thought is ours and that the 'I' in that thought can be intersubstituted with the 'I' in this one, as for

example we apprehend and apply criteria of identity for days in judging that a 'today'-occurrence on one occasion is intersubstitutable with a 'today'-occurrence on another. But this is not the way that 'I' works. It is the singular property of 'I' that no criteria of identity for selves play a role in reidentification of the self across different thoughts, or in intersubstitution of the 'I' in one thought for the 'I' in another. There is no judgment of identity across contexts of the kind we need to make with 'here', 'today', or 'this chair', and there is no risk of mistake of the sort we can always make in those cases. We can mistakenly intersubstitute a pair of 'here'-occurrences that pick out different places and we can mistakenly intersubstitute a pair of 'today'-occurrences that fall on different days, but it's hard to make sense of the idea of mistakenly intersubstituting someone else's 'I'-occurrence for one's own. From a psychological point of view, the bonds between our thoughts and the boundaries that separate them from those of others are fundamental and unanalyzable.

There is a general recognition in the literature on the self that the self isn't picked out by an individuating conception of the kind that could serve as a Fregean sense. Shoemaker has emphasized this feature of 'I' in his own work:

> My use of the word 'I' as the subject of [statements such as 'I feel pain' or 'I see a canary'] is not due to my having identified as myself something [otherwise recognized] of which I know, or believe, or wish to say, that the predicate of my statement applies to it. (Shoemaker 2001)

It has also been discussed also by Strawson, Evans, Anscombe, and others, and is often expressed by saying that no criterion of identity is employed in picking out the self as the object of a referential act. Kant recognized the fact clearly:

> In attaching 'I' to our thoughts, we designate the subject . . . without noting in it any quality whatsoever—in fact, without knowing anything of it either directly or by inference. (Kant 1997, A355)

What these passages, and others like them in the literature, don't express clearly enough to my mind is the importance of the fact that criteria of identity are not employed in *re*identification. It's the lack of employment of criteria of identity in judgments of same-self relatedness across contexts that distinguishes self-identification from other kinds of indexical and demonstrative reference.

All of this can seem to lead inevitably to the conclusion that selves are primitive constituents of reality, mutually impenetrable substances in which thoughts are housed. As James puts it:

> Each [self] keeps its own thoughts to itself. There is no giving or bartering between them. No thought even comes into direct sight of a thought in another [self] than its own. Absolute insulation, irreducible pluralism is the law. . . . Neither contemporaneity, nor proximity in space, nor similarity of quality and content are able to fuse thoughts together which are sundered by this barrier of belonging to different personal minds. The breaches between such thoughts are the most absolute breaches in nature. (James 1990)

'I' works just like other indexicals on this view. The reason we can't say anything informative about what makes a pair of thoughts or impressions thoughts or impressions of the *same* self is that there is nothing more informative to say. Selves are basic constituents of reality, and criteria of identity for selves can't be analyzed in any other terms. This view is probably closest to that of the man on the street. Ask a nonphilosopher what makes two occurrences of 'I' coreferential and he will tell you they belong to the same self, end of story.

That there is something more interesting going on here is indicated by several things. First, it makes possession of the notion of a self prior to the ability to reidentify oneself over time, and that doesn't seem phenomenologically right. It seems too much thought to attribute to the child, for example, who says I enjoyed the ice cream yesterday and I want to have some more. The ability to reidentify oneself over time seems to precede any grasp of the concept of a self. More importantly, it suffers from problems that beset any view that requires intersubstitution to be regulated by *judgment*s of same-self-relatedness. Any judgment that spans a pair of temporally separated thoughts would seem to presuppose, and cannot therefore establish, the numerical identity of the judging self. Instead of pressing these objections, which would require looking quite deeply into the nature of judgment, I want to suggest an alternative account, starting with an example drawn from Perry (1986) and embellishing it in a way that will allow us to see how to derive the eccentricities of 'I' from a story about the mechanics of achieving reference.

Preconceptual Thought *About* and Post Hoc Conceptualization

Perry (1986) considers a fictional populace who live in a place called Z-land. Z-landers are unadventurous folk who never travel and who never meet, or have occasion to talk to, people from other places. They talk about the weather, and their weather beliefs play an important role in their practical lives; what they plan and how they act depends on their beliefs about the weather.[7] What is striking about their talk about the weather is that it

never includes an explicit reference to a place. They never mention or make any reference to Z-land in their reports; the nightly news says simply "rain today, sun tomorrow, storms on Wednesday" without any mention of place, and that tells them everything they need to know in order to decide whether to cancel the picnic or carry an umbrella.

Reports for Z-landers serve as expressions of belief and we assume that they reflect the structure of those beliefs in containing the same semantically significant constituents. We should agree with Perry that the Z-landers' beliefs are about the weather in *Z-land*, even though they make no mention of place; Z-land appears in the specification of their truth conditions. It is the weather in Z-land that confirms or disconfirms them. Perry puts this by saying that Z-land is an *unarticulated constituent* of their weather beliefs. The Z-landers can make the spatial content of their weather talk explicit by adding 'here' to their reports. This is a practically trivial addition so long as they remain at home; 'here' can function grammatically as a name for Z-land, and unconstrained intersubstitution of tokens of it will preserve reference. Inferences will go through unproblematically, information will flow smoothly from Z-land to Z-lander weather beliefs and from one Z-lander to another in talk about the weather, and the sheer mechanics of the situation will ensure coordination without any need for a mediating concept of the place. If Z-landers begin to move, however, or communicate with non-Z-landers, unconstrained intersubstitution will lead to faulty inferences and they will need something in the head that can serve as a constituent of beliefs about the weather there and distinguish those beliefs from beliefs about weather elsewhere. They will need, in short, an idea of Z-land that occurs as an articulated constituent in beliefs that are sensitive to the state of, and guide behavior in, Z-land. The links between perception, belief, and action will be less direct than they were before, mediated now by a representation of place and self-ascription of location. Z-landers then will not respond immediately to reports of rain by grabbing an umbrella; they will ask *where* it is raining and they will stop to check it against the place they are *at*.

For another example of the process by which unarticulated constituents are articulated, think of the child who learns to use 'mom' and 'dad', 'mother' and 'father' to refer to his own mother and father, without understanding that motherhood and fatherhood are relations.[8] So long as he remains in the family circle, this isn't something he needs to worry about.[9] If he hears that mom is in the garden, he will know whom to find there. And if someone tells him that mom is wearing a straw hat, he will conclude correctly that there is someone in the garden who is wearing a straw hat.

'Mom'-occurrences in his restricted world will all lead to the same place, and there is no need to interpose a parameter to coordinate them either with one another, or with their common referent. For him, 'mom' functions as a singular term. As soon as he leaves the family circle, however, he will discover that 'mother' refers to other people in the mouths of his friends, he will need to keep explicit track of *who* is speaking, and he will need to know something more about motherhood. *What* he needs to know, precisely, is a complex and highly context-dependent question. The general rule is that he needs a way of getting from an utterance of 'mother' to its referent and a way of determining whether a pair of 'mother'-occurrences corefer. What this requires will depend on contingencies of his situation, for example, how many mother–son pairs he knows, and whether there are general ways of identifying speakers' mothers.[10] How much structure he needs in his own head to keep the information flowing smoothly so that information about a particular mother is collected, combined, kept safely separate from information about others, and brought to bear in the right way in his interaction with her, will depend on which other mothers he gets information about, and how. Mom-talk in the child's environment, and mom-thought in the child's head, while he is confined to the family circle, like Z-lander weather reports, have a hidden contextual parameter[11] that can remain hidden so long as he remains in a context in which it has a fixed value but that needs to be articulated when he is exposed to contexts across which its value varies.

Focus on the process of replacing unarticulated constituents with representations. Recall how things went with the Z-landers. They began by talking freely of the weather without overt acknowledgment of the spatial relativization of weather talk. Making the spatial relativization explicit proceeds in two stages; they learn first to attach a 'here' to their weather reports as a caution against unconstrained intersubstitution. The 'here' signaled to them that they shouldn't infer disagreement from differences in reports, for the reporters might live in different places. But before they can begin to use weather talk again in inferences, before their weather thoughts can be properly integrated into the web of belief, and before they can coordinate their own weather reports with weather reports of their neighbors, they will need to replace 'here' with a parameter that explicitly represents, by varying in value with, their own location. Before they can be used effectively in reasoning, that is to say, weather reports and weather beliefs that concern different places will have to be made *internally* distinguishable. Indexicals are a perfectly legitimate way of securing reference and they have the critical advantage of securing it in a way that bypasses the

ordinary route through ideas, placing less conceptual burden on their users. You don't have to have an individuating conception of a place to refer to it indexically. But they have the disadvantage of concealing relations of intersubstitutivity. One cannot tell from an internal perspective—by inspection, so to speak, of the beliefs themselves—whether a pair of 'here', 'this', or 'today' occurrences are intersubstitutable. From the inside, one 'here' or 'today' thought looks like any other.

This second stage, at which the indexical is replaced with explicit representation, is highly nontrivial. To pull it off, Z-landers will have to come to appreciate and employ criteria of identity for places. They will have to learn to identify and individuate places, and they will need a general concept of places that are not their own, something that they didn't need in order to refer to them successfully using 'here'. The same story could be told about sheltered children. Once they have left the family circle and can no longer rely on the context to coordinate reference, if they are to use 'mother' effectively in reasoning, they will need to develop an understanding of familial relations and learn to apply criteria of identity in deciding whether distinct 'mother'-occurrences are intersubstitutable. As soon as they begin to move around, Z-landers no longer have the kind of built-in, invariant informational link to a place that allows it to act as an unarticulated constituent in their thought and they need something a little higher up in the representational hierarchy, something with a looser connection to experience, to hang their Z-land thoughts on.

Return now to bridging the gap between Descartes's starting point and reference to a temporal continuant. Recall the terms of the problem. He started with an implicit definition of the object of inquiry as that thing, whatever it is, whose existence is made known in the production of an 'I'-thought. This prohibits him from appealing to any constraints on the nature of selves but those that are conditions of the possibility of 'I'-use. I want to suggest that we learn to use 'I' by a mechanical procedure: producing tokens, and allowing unconstrained intersubstitution, relying on the external relations among the tokens to preserve reference. Reference is thus secured without explicit representation of criteria of identity—indeed, without invoking concepts of any kind. This reverses the ordinary direction of determination. We do not introduce the concept of a self governing intersubstitution of 'I'-tokens; it is the *de facto* relations among tokens that get intersubstituted by application of the mechanical procedure that determines the criteria of identity for selves. It is thus, and only thus, that we can refer to ourselves not only without having an individuating idea of our own selves, but without knowing what selves *are*, without

knowing what makes two thoughts thoughts of the *same* self. This makes it wholly unique among indexicals. In all other cases, the rules of use for the indexical invoke criteria of identity applied in reidentification and intersubstitution. Competence with the terms requires apprehension of such criteria. You have to possess criteria of identity for days and mothers to sort 'today'- and 'mom'-occurrences into intersubstitutability classes. No conceptualization of this sort is provided by the rule of use for 'I'.

Many in the contemporary literature have emphasized that conceptualization is, however, essential to being properly ascribable a genuine idea of self. Evans, for example, holds that in order for 'I' to be recognized as a semantically significant component of thought,

> one's Idea of oneself must also comprise, over and above [unmediated links to perception and action], a knowledge of what it would be for an identity of the form 'I = ?' to be true, where ? is . . . an identification of a person which—unlike one's 'I'-identification—is of a kind which could be available to someone else. . . . (Evans 1982, 209)[12]

Kant made the same point in distinguishing the 'I think' that is the empty accompaniment of every thought from meaningful self-ascription. For Evans, the requirement is an instance of a general condition on identifying thought that gets its justification from an understanding of the role of thought in mediating perception and action. He writes:

> sensory input is not only connected to behavioral disposition in the way I have been describing . . . but also serves as the input to a thinking, concept-applying, and reasoning system. . . . it is only those links which enable us to ascribe content (conceptual content now) to the thoughts.[13]

He's absolutely right about that. Reasoning with a term requires a practical understanding of how it interacts inferentially with other terms. To represent oneself in the sense of being able to take oneself as an *object* of thought requires more than the ability to produce mental tokens of 'I'. Those tokens have to function as semantically significant constituents; they have to be connected in an inferentially articulated network of concepts. We have to be able to associate them with other singular terms, use them in inferences, and distinguish them from other objects and from one another. We have to be able to entertain thoughts not only about our own selves, but those of others. We have to know, in sum, what would make true an identity of the form 'I = *g*', where '*g*' is an arbitrary name or identifying description.

The repeated pattern is the one that we saw with the Z-landers: there is a suppressed parameter with a fixed value that is first expressed indexically and later replaced with a referring term. We get a closer analogy if we

embellish. Instead of Z-landers, imagine a population of conscious, weather-sensing trees: Tree-landers. Tree-landers can tell sunny weather from rain, and each has a repertory of actions suited to weather in its own locale. Perhaps, for example, they open their leaves in rainy weather and close them in sun. They get information about rain in other locales only indirectly via reports from birds that bring news from distant lands. Their weather thoughts and reports start out, like the Z-landers, without explicit spatial content. They learn to attach 'here' mechanically, allowing unconstrained internal intersubstitution but refraining from substitution in connection with reports from outside. The conceptualization needed to integrate their here-thoughts properly into the web of belief occurs at a later stage, if it occurs, and involves apprehension of objective criteria of identity for places.

In apprehending criteria of identity for selves, we face the something like the task that Tree-landers face in apprehending criteria of identity for places. Ideas of other locations are for them, in the beginning, slots where they funnel information about weather that is not weather *here*, that is to say, weather that doesn't make itself felt in *this* way and have a direct, regulative bearing on the opening and closing of leaves. And just so, ideas of other *selves* are, in the beginning, for each of *us*, places where we house thoughts and experiences that are not our own, that is to say, thoughts and experiences other than *these,* thoughts and experiences that don't make themselves felt in this way and have a direct regulative bearing on activity. They are whatever lies at the other end of the linguistic chains that bring us news of such thoughts. There is nothing inherently mysterious, if we focus in this way on the sheer mechanics of securing reference, in the fact that each of us can refer to one object in thought in a way that is immune to failure of reference (i.e., whose existence is ensured and reference to which is secured in the very production of the thought), which is nevertheless distinct from (because it possibly outlives) the thought, and yet with no conception of what kind of thing we are referring to.

This account gives us:

1. Identification without application of criteria of identity
2. Immunity to failure of reference
3. Reference to a continuant

Moreover, since no criteria of identity are employed in reidentification, apprehension of criteria of identity, if and when it is made, will present itself as a *discovery* and show all of the psychological signs of contingency. Here's what I mean by saying it will show the psychological signs of contingency. Suppose *R* is the relation that turns out to unite the thoughts and experi-

ences of a self, which is to say, the relation that 'I'-occurrences which get intersubstituted by application of the mechanical procedure bear to one another. It doesn't matter what *R* is; so long as it doesn't make explicit and ineliminable reference to a self, the proposition '*S* and *S** belong to the same self iff *S R*s *S**' will not look like a conceptual truth.

1. It will be dubitable.
2. Being told that *S R*s your present thoughts will seem to leave it as an open question whether *S* is nevertheless *yours*.
3. It will seem as though you can imagine scenarios in which all of the objective *R* relations remain fixed but the gaze of your consciousness crosscuts them, alighting now here, now there, at will.

These facts, when they are stated in general form, will seem to rule out any reductive candidate for a constitutive relation between temporal parts of a self. All of them have their roots in the fact that judgments of ownership and co-ownership of mental states over time are not mediated in the ordinary way by application of concepts. This is what makes 'I' wholly unique, even among indexicals. We might say that in this case, unlike other indexicals, we are constitutively confined to a context in which intersubstitution preserves reference. (1)–(3) have provided the basis for explicit philosophical arguments against reductive analyses and they are unquestionably implicated in the intuitive conception of the self as a primitive locus of mental life, a metaphysical free variable only contingently connected with the material contents of the universe. As Nagel has said:

> the very bareness and apparent completeness of the concept [of the self] leaves no room for the discovery that it refers to something that has other essential features which would figure in a richer account of what I really am. Identification of myself with an objectively persisting thing of whatever kind seems to be excluded in advance. (Nagel 1989, 35)[14]

And earlier:

> My nature . . . appears to be at least conceptually independent not only of bodily continuity but also of all other subjective mental conditions, such as memory and psychological similarity. . . . at the same time it seems to be something determinate and nonconventional. That is, the question with regard to any future experience, "will it be mine or not?" seems to require a definite yes or no answer. . . . (Ibid., 34)

> This seems to leave us with the conclusion that being mine is an irreducible, unanalyzable characteristic of all my mental states, and that it has no essential connection with anything in the objective order or any connection among those states over time. (Evans 1982, 34)

I am suggesting that there is a mundane explanation that avoids this conclusion.

Drawing the Boundaries around the Self

So let's return to the question of what makes one I-thought fall within the referential scope of another on the assumption that no criteria of identity are employed by 'I'-users in reidentification. Or, to put it in a way that evokes Sartre, what makes this the case on the assumption that the self is constituted in, rather than presupposed by, the act of reflective self-consciousness? Anscombe thought that Descartes's argument foundered on what she calls the "intolerable difficulty" that if the thinker outlives the thought—that is, if the 'I' in one thought is even potentially intersubstitutable with the 'I' in others—its existence can't be revealed in the mere production of an I-thought. Her worry was that the notion of an I, in a sense that allows *different* occurrences of 'I' to refer to the *same* self, must be prior to and independent of—and cannot be *constituted* by—the mere production of those thoughts.

We have seen now why Anscombe's difficulty is not after all so intolerable. There are two ways to define an intersubstitutability context: implicitly and explicitly. Explicit definition does indeed require prior conceptualization of the referent, but if we can learn how to use 'I' by a purely mechanical procedure, one that can be described without any explicit reference to or conceptualization of its target, we can let the de facto relations among tokens that in fact get intersubstituted when 'I' is used in accordance with the procedure implicitly define the referent. Reference—even reference to a continuant—can in this way *precede* apprehension of criteria of identity. Apprehension of such criteria will be unnecessary for purposes of coordination and if it is ever made, it will present itself as a discovery. We don't need semantic links between terms (or their mental analogues, ideas) and their objects where there are architectural links in place. It is only when we break architectural links by moving to different contexts that we require mediating concepts. That was the insight in Perry's original example, but it has been extended here to apply to the links between the parts of the self in a way that explains some of its most intractable peculiarities. It shows how it can turn out to be news to us that the thing we've been identifying and reidentifying all along is really one whose parts are, unbeknownst until now, held together by such and such connections. If we specify the rules of use ("attach 'I' indiscriminately to all thought and experience, allow unconstrained intersubstitution") and let the criteria of identity be

implicitly defined by the rules of use, then finding out what selves are, and in particular, what *I* am, will be a matter of determining what turns out to unite occurrences that get intersubstituted by the procedure specified in the rule of use. The crucial insight here is that we don't need to represent those relations in order to reidentify the self.

What do we get if we apply this strategy? We get—I propose—external informational relations, the direct causal links inside the head that permit information to flow without passing through perceptual or linguistic channels. These are what tie the events that lie along a single stream of consciousness to one another and separate them from events in the psychological histories of others. Selves are nothing more than sealed pockets of world-representing structure, communicating with one another and with the environment through controlled channels, mediated by experience. Each of us applies 'I' mechanically by attaching it as grammatical subject to internal states and only later, once we have worked up a rudimentary picture of the world, do we raise questions about its own nature. It is really the existence of these sealed pockets—Humean bundles, connected by external relations that let information flow directly between them—that provide the contexts of intelligibility for 'I'-use. The intelligibility of 'I'-use doesn't depend on the existence of simple enduring substances for thoughts to inhere in; it requires nothing more than the various I-thoughts, the experientially unmediated informational relations they bear one another, and ordinarily enough internal unity and informational segregation from the environment to make 'I'-use relatively unproblematic.

It is crucial to this way of understanding things that the kind of connectedness in question is an external relation, one that doesn't supervene on the internal character of such states and that is indiscernible from the inside. It is true that 'I' can meaningfully occur only in the context of articulated world-representing structure, that this requires a certain amount of continuity of content, and that given certain assumptions about how world-representing structure is formed, that in its turn may require continuity in qualitative character. And it is also true that causal connectedness of the sort that permits experientially unmediated flow of information ordinarily has the *effect* of imposing a kind of continuity in content, but internal continuity, if I am right, plays no role in individuating selves, and—again—is in no way constitutive of the relation *belongs-to-the-same-self-as*. We might want to say that internal criteria are important and relevant to *personhood*, for that is a social and moral status. The notion of a person is entangled with concepts of agency and responsibility. The

existence of persons, in this richer sense, is not guaranteed by the production of an I-thought, but only external criteria are constitutive of the identity of Cartesian *selves*.

This brings out a critical divide among forms of psychological reductionism often overlooked in the literature on personal identity. If we call views that share this feature pure Lockean views, and views that include internal criteria among the conditions of selfhood, Parfitean, after their most conspicuous exemplars, Locke and Velleman hold Lockean views. As does James, who writes:

> My present Thought stands . . . in the plenitude of ownership of the train of my past selves, is owner not only de facto, but de jure, the most real owner there can be. Successive thinkers, numerically distinct, but all aware of the past in the same way, form an adequate vehicle for all the experience of personal unity and sameness which we actually have.[15]

According to James, the self or subject

> is a Thought [mental state], at each moment different from that of the last moment, but appropriative of the latter, together with all that the latter called its own. All the experiential facts find their place in this description, unencumbered with any hypothesis save that of the existence of passing thoughts or states of mind.[16]

As a theory of the self, only a pure Lockean view—that is, one that allows no admixture of internal criteria into the constitutive relations between the parts of a self—can explain how we can identify and reidentify ourselves over time without application of any criteria whatsoever. Only the pure Lockean view explains the fact, wholly unique to self-identification and critical to Descartes's argument, that we can make repeated identifying reference to ourselves in thought without having (or needing to have) any conception of what unites our temporal parts.

What unites the states of a subject, on a pure Lockean view, are external relations that allow for perceptually unmediated communication. This is so both synchronically and diachronically. I have perceptually unmediated access to my own past and present thoughts and experiences. The only access I have to those of others is mediated by the production of publicly observable symbols or behavior. They have to talk, write, or behave in ways that I can see, ways that are designed to clue me in. Although there is no contradiction in supposing these kinds of consciously unmediated informational connections cutting across the boundaries of bodies, in fact, the body provides the physical context in which they are realized, maintaining the direct channels for the propagation of information and sealing them off from the external environment in a way that forces infor-

mation from the outside to be funneled through experience. The only way you can work your thoughts into mine is by the production of observable behavior or information-bearing symbols. My own thoughts, by contrast, feed directly into one another. There are immediate connections among the mental states of a single subject, but immediacy in this sense is shorn of the traditional encumbrances of transparency and certainty, and compatible with the kind of complicated action behind the scenes (processing that is invisible to the subject, processing that doesn't get represented *in thought*) that Dennett emphasizes in his critiques of the Cartesian Theater, and importantly, it can accommodate the kinds of pathologies and internal divisions of self that we know to arise when psychological boundaries inhibit the flow of information.[17]

There is room in this for both indeterminacy and legislation. How indirect and artificial can the causal relations in question be? If there is a difference between recovering memories of one's own and acquiring someone else's by artificial means, there have to constraints on the causal process by which I come to have them, that is to say, the route by which they come to be represented in my thought. Obvious counterexamples will tell us what kinds of constraints need to be added, and borderline cases will simply be evidence of vagueness in the concept. It might be that under certain conditions—for example, if fission or amnesia becomes common enough, or memories can be bought and uploaded with no internal signs of inauthenticity—the notion of a temporally extended self, as we know it, will no longer have application. We will need a more articulated set of concepts that accommodates the full range of relations one can bear to downstream descendents of one's thoughts.

What is special about 'I', wholly unique to it as a term in thought, is that the facts that govern intersubstitutability of 'I'-occurrences are external to, and need not be represented in, thought. To employ a Wittgensteinian image, we don't represent our boundaries so much as bump up against them. I don't have to know what I am to know which thoughts, memories, experiences, hopes, are mine. I don't have to know anything *about* my future self, that is, the downstream descendents of my present thoughts, experiences, and so on, to have thoughts about, make decisions on behalf of, have hopes and fears, and so on, centered on *her*. Whatever ideas I have about my future self play no role in determining which of the world's future inhabitants is me. If I think that I will be the first woman president of the United States, but Marla Maples beats me to it, that doesn't make my present thoughts about my future self thoughts about Miss Maples. Whoever gathered these memories and sent them to me along the

discriminating internal channels that Nature has provided inside the head, with whatever internal embellishments and emotional resonances that past self added, that was me. And whoever receives these thoughts and experiences along the same channels will be me as well. Am I a thing? I am a thing in the same way that a river is a thing. When I talk or think about myself, I talk or think about the connected, and more or less continuous, stream of mental life that includes this thought, expressing the tacit confidence that that is a uniquely identifying description (in the same way I might speak confidently of 'this river' pointing at a part of it, expressing the tacit assumption that it doesn't branch), but it need not be. *Belonging-to-the-same-subject-as*, in this sense, is not an equivalence relation, and if that means that subjects are not objects, then there is no object in the world to which the Cartesian 'I' can be said to refer. But that doesn't mean there are no selves.

To sum up, we have a term in the language of thought that functions like a singular term, introduced by the mechanical procedure 'attach indiscriminately to mental states, intersubstitute freely'. And if we let the procedure implicitly determine criteria of identity, we end up saying that particular tokens of it refer to connected streams of world-representing structure in which they are situated. This positive view of identity over time is not new. It has a long history and has been eloquently and ably defended in recent literature by Velleman.[18] The story told here derives it from a general semantic framework in which thought about the self can be situated alongside other indexicals together with thought that conforms more closely to the Fregean model. The framework develops Perry's notion of an unarticulated constituent in a different direction than it has been taken in the literature,[19] but closer to the spirit of the original proposal. Unlike Perry's own explicit account of thought about the self, however, which makes the notion of a person a constituent of the self-concept, it allows reference to *precede* conceptualization. That is crucial to understanding the peculiar epistemic predicament that 'I' presents, so vividly rendered in the appended quotes from Nagel and James, and exploited by Descartes in the argument of the Second Meditation.

We can say quite precisely where that argument goes wrong. Descartes began with the question: What is this thing identified in the mere thinking of an I-thought and (here making explicit what he tacitly assumed and what Anscombe argued he was not entitled to) *re*identified with every reflective thought and experience? The argument departed from the observation that we have a way of referring to ourselves that is guaranteed reference and that leaves the nature of the self entirely open. And it turned on

the thesis that nothing can be true of one's self that is not made known in the act of thinking (premise 2 in the formulation above):

(2) I can't be identical to anything whose existence can be doubted.

In Augustine's version of the argument, (2) takes the form of the claim that the mind "is certain of being that alone, which alone it is certain of being" (Augustine 2002, book X). The Tree-landers gave us a concrete understanding of how it *could* be possible to form thoughts about oneself without having any conception of what kind of thing one is, that is, how we manage to refer in thought to an object of whose existence we are assured, and *know* that we are assured, but of whose nature we are *entirely* and *completely* ignorant. And all of this without the self being anything but a perfectly mundane sort of thing.

The recognition of unarticulated constituents is a deep and consequential emendation of a Fregean view of reference, but I believe it is correct, and that nothing less is needed to get the right account of thought about the self. Descartes's mistake is to go from the *epistemic* basicness of judgments of same-self identity in the first-personal case to ontological basicness, that is, to elevate the fragile boundaries that separate selves—the contingencies that support reidentification without application of criteria of identity over time—to the status of "absolute breaches in nature" (James 1950).[20]

Notes

1. "Now it is very certain that this notion, thus precisely understood, does not depend on things whose existence is not yet known to me" (Descartes 1984, 103).

2. Nagel is expressing, but not endorsing, the view in this passage.

3. I'll be sidestepping most exegetical questions, assuming a fairly standard reading of the argument of the *Second Meditation* and interpreting Anscombe's remarks quite freely.

4. Natural languages are, of course, massively ambiguous, and wholesale eradication of ambiguity may not be practically possible. Still, each particular instance of ambiguity is in principle eliminable.

5. In the best case, you would have an understanding of what Evans called the fundamental ground of difference between places, criteria for individuating places across all possible worlds. For practical purposes, however, a relatively reliable guide for reidentification that provides a basis for substitution suffices. What counts as reliable enough, and which circumstances reliability is relative to, will depend on what the practical purposes are.

6. There can be vagueness. 'Here' and 'now', for example, are notoriously vague, even when we take account of all potentially disambiguating features of conversational context. When I say to you in Sydney "Let's meet back here in an hour," we both probably know that Calcutta isn't included in the extension, but there is probably no determinate fact about which exactly of the indefinitely many surveyable regions of space centered on our location I referred to. It is not that reference must be determinate, but that it is not *more* determinate than the boundaries supplied by any explicitly or tacitly provided sortal.

7. We suppose that the weather is uniform across Z-land.

8. In formal terms, the process of articulating unarticulated constituents is that of representing ones frame of reference. One makes concepts drawn from different frames of reference communicable by plotting them jointly in a frame that includes a dimension in which their relations can be made explicit. The new constituents are values of parameters in added dimensions.

9. Supposing that he doesn't get information from other sources (books, radio, television), or at least that those sources don't use the terms 'mom' and 'dad'.

10. If sons never leave their mothers' sides or if sons were always the spitting images of their mothers, the task would be easier than in a world like ours, in which there are no foolproof visual ways of recognizing mothers.

11. Not hidden in the sense that it corresponds to anything in his head, but in the sense just described: it plays an unacknowledged role in coordinating his thought with its intentional object.

12. Conceptualizing oneself, for Evans, means conceiving of oneself as one among a range of entities, any of whom can bear the properties that one ascribes oneself immediately on the basis of experience.

13. The condition, which he calls the Generality Constraint, formulates the conditions under which an internal symbol makes a contribution to content, and can be said to stand for, or *represent*, a thing. In its general form, the condition requires us to see "the thought that *a* is *F* as lying at the intersection of two series of thoughts: the thoughts that *a* is *F*, that *a* is *G*, that *a* is *H*, . . ., on the one hand, and the thoughts that *a* is *F*, that *b* is *F*, that *c* is *F*, . . ., on the other" (Evans 1982, 209).

14. Nagel is not endorsing the view he is articulating here.

15. See James 1950, 1.360, and James 1990.

16. See James 1950, 1.400–401, and James 1990. See also James 1950, 1.338–342.

17. For an account of immediacy also shorn of those traditional encumbrances, see Moran 2001.

18. See "Self to Self" and "Identity and Identification" in Velleman 2005. I owe much to Velleman's masterful exposition of the view.

19. See, e.g., Recanati 2001 and Stanley 2002.

20. The themes in this essay are also developed in my book, Ismael 2006.

References

Anscombe, G. E. M. 1990. The First Person. In *Mind and Cognition*, ed. W. Lycan. Oxford: Blackwell.

Augustine. 2002. *On the Trinity*. Ed. G. B. Matthews, trans. S. McKenna. Cambridge: Cambridge University Press.

Descartes, R. 1984. Meditations on First Philosophy. In *The Philosophical Writings of Descartes*, vol. II. Ed. and trans. J. Cottingham, R. Stoothoff, and D. Murdoch. Cambridge: Cambridge University Press.

Evans, G. 1982. *Varieties of Reference*. Oxford: Oxford University Press.

Ismael, J. 2006. *The Situated Self*. Oxford: Oxford University Press.

James, W. 1950. *The Principles of Psychology*, vol. 1. New York: Dover.

James, W. 1990. The Stream of Consciousness. In *Mind and Cognition*, ed. W. Lycan. Oxford: Blackwell.

Kant, I. 1997. *Critique of Pure Reason*. Trans. P. Guyer and A. Wood. Cambridge: Cambridge University Press.

Kaplan, D. 1989a. Demonstratives. In *Themes from Kaplan*, ed. J. Almog, J. Perry, and H. Wettstein. Oxford: Oxford University Press.

Kaplan, D. 1989b. Afterthoughts. In *Themes from Kaplan*, ed. J. Almog, J. Perry, and H. Wettstein. Oxford: Oxford University Press.

Moran, R. 2001. *Authority and Estrangement*. Princeton: Princeton University Press.

Nagel, T. 1989. *The View from Nowhere*. Oxford: Oxford University Press.

Perry, J. 1986. Thought without Representation. *Supplemental Proceedings of the Aristotelian Society* 60:263–283.

Recanati, F. 2001. Literal/Nonliteral. *Midwest Studies in Philosophy* 25:264–274.

Shoemaker, S. 2001. Self-Reference and Self-Awareness. In *Self-Reference and Self-Awareness*, ed. A. Brook and R. C. DeVidi. Amsterdam: John Benjamins.

Stanley, J. 2002. Making it Articulated. *Mind and Language* 17:149–168.

Velleman, D. 2005. *Self to Self*. Cambridge: Cambridge University Press.

MIT Press books may be purchased at special quantity discounts for business or sales promotional use. For information, please email special_sales@mitpress.mit.edu or write to Special Sales Department, The MIT Press, 55 Hayward Street, Cambridge, MA 02142.

This book was set in Stone and Stone Sans by Westchester Book Group. Printed and bound in the United States of America.

Library of Congress Cataloging-in-Publication Data

Time and identity / edited by Joseph Keim Campbell, Michael O'Rourke, and Harry S. Silverstein.
p. cm.—(Topics in contemporary philosophy)
"A Bradford book."
Includes bibliographical references and index.
ISBN 978-0-262-01409-0 (hardcover : alk. paper)—ISBN 978-0-262-51397-5 (pbk. : alk. paper)
1. Time. 2. Identity (Philosophical concept). 3. Self (Philosophy). 4. Death.
I. Campbell, Joseph Keim, 1958–. II. O'Rourke, Michael, 1963–. III. Silverstein, Harry, 1942–.
BD638.T5437 2010
115—dc22

2009041935

10 9 8 7 6 5 4 3 2 1