Framing the Problems of Time and Identity

MATTHEW H. SLATER

CONCEPTS FAMILIAR YET PERPLEXING

Many philosophical concepts are difficult. Some, however, are doubly confounding in their apparent familiarity. The concepts of Time and Identity may top this short list. The fourth century philosopher, Augustine of Hippo, expressed his exasperation with time this way:

For what is time? Who can easily and briefly explain it? Who can even comprehend it in thought or put the answer into words? Yet is it not true that in conversation we refer to nothing more familiarly or knowingly than time? And surely we understand it when we speak of it; we understand it also when we hear another speak of it. What, then, is time? If no one asks me, I know what it is. If I wish to explain it to him who asks me, I do not know. (Confessions, 11, XIV, 17)

Time is an intricate part of our experience. Indeed, it can seem like the passage of time is a prerequisite for having any experiences at all. But a number of recalcitrant questions about time remain: Is time real? Does time objectively flow? Do past and future moments exist or is the present in some way special?

The concept of identity also seems at once philosophically unproblematic and frustratingly difficult. It is, as philosophers have long noted, that relation everything bears to itself and to no other thing. Of course, as Hawthorne (2003, 99) points out, this cannot be an analysis of identity — for ‘itself’ and ‘no other thing’ already presuppose an understanding of identity. David Lewis likewise tempers his praise of identity’s simplicity:

There is never any problem about what makes something identical to itself; nothing can ever fail to be. And there is never any problem about what makes two things identical; two things never can be identical. There might be a problem about how to define identity to someone sufficiently lacking in conceptual resources — we note that it won’t suffice to teach him certain rules of inference — but since such unfortunates are rare, even among philosophers, we needn’t worry much if their condition is incurable. (Lewis 1986, 192–193)

And yet philosophers face plenty of hard questions about identity. Is a statue identical with the bronze out of which it is made? Is identity ever contingent or indeterminate? How can things change and yet remain the same? What determines whether a person is the same over time — do we have privileged first-person access to the sameness of “the self”? What, in general, are the metaphysics and epistemology of persistence?

Many of these questions (explicitly or implicitly) involve time. They thus tend to inherit its complexities! What we say about persistence presumably bears on what we say about time (and vice versa). The essays in this collection attempt to make progress on some of these most baffling philosophical problems, either singly or in combination. Before getting to them, however, let us sketch some of terrain lying in the background.

TIME

The Reality of Time

Perhaps the best place to begin a discussion of the philosophy of time is with the question of whether time exists or is real. The ancients seemed to have presupposed the reality of time, even if it wasn’t always clear what to say

---

1 See, for example, C. S. Peirce (1982, 455).
about time — whether the “flow” of time depended on other things, whether the future is knowable or fixed, whether time is eternal, and so on).\(^2\) It wasn’t until the twentieth century — ironically, when physicists seem to gain a better scientific understanding of time — that philosophers began to question whether time was in fact real. One argument in particular (McTaggart 1908) has influenced subsequent discussion. Let's begin there, as it will also serve to introduce some of the complex and overlapping terminology employed by various philosophers of time.

McTaggart’s argument began with a basic distinction between two ways of ordering events in time. The first — what he called “the A-series” — locates events relative to the present. For example, my birth is in the past and death is in the (hopefully somewhat distant) future. But these temporal properties change: my birth was once in the future and my death will someday be in the past. The second way of ordering events in time — what McTaggart called “the B-series” — remains constant. My birth is earlier than my death but later than that death of Socrates. The temporal relations remain constant: it does not matter when the proposition ordering my birth and death is uttered. Its truth-value will never change, for it concerns only the temporal ordering of events. We might say that A-series expressions (’tomorrow’, ’now’, ’yesterday’, &c.) are “tensed” expressions while B-series expressions (’earlier than’, ’later than’, ’at 1PM PST, November 25th, 1977’, &c.) are “untensed” (or “tensless”) expressions.

So far, so good. But here McTaggart points out that if time is real, there must be an A-series — for A-series expressions are more fundamental or more characteristic of “temporality”. But unfortunately, thought McTaggart, the A-series is contradictory and thus cannot exist. As time flows, events come to have different A-series properties. But the propositions that ‘My birth is past’ and ‘My birth is future’ are incompatible. Thus, because the A-series does not exist, time is unreal. Philosophers have gone several different ways responding to this argument. So-called “B-theorists” (sometimes called “Eternalists”) object to McTaggart’s assumption that B-series facts do not afford “real time”. “A-theorists” (sometimes called “Presentists”) object to his claim that the A-series is contradictory. While these philosophers agree that time is real, they propose (purportedly) different metaphysical accounts of time.\(^4\)

**Eternalism & Presentism**

We might identify two reasons for thinking with McTaggart that the B-series alone cannot be an account of time. First, we might regard the A-series as semantically fundamental — if the B-series ordering depends on A-series properties, and those properties turn out to be incoherent, then follows that there just is no B-series. Second, the B-theory can appear to be problematically “static”, but time is dynamic. The present “flows” ineluctably. But might claim that though time cannot properly be said to “pass”, we can still make sense of A-series locutions as indexicals. Here the analogy between ‘here’ and ‘now’ becomes useful. Imagine standing in the middle of a long stretch of desert highway. “It’s hot here,” you whimper. Your philosopher friend tries to placate your fried nerves: “That’s false! There is no spot which is objectively ‘here’!” He’s misguided, though (in several ways). Your statement can be true from where your standing, even if it’s false in Siberia.

---

\(^2\) As with much of philosophy Aristotle has some particularly rich discussions of time; see in particular his *Physics* (Book VIII) and *De Interpretatione* (Chapter 9).

\(^3\) Though as we shall see below, it may not be quite right to identify these positions.

\(^4\) For an excellent discussion of McTaggart’s argument and responses to it, see Le Poidevin (2003, 127–147).
This is roughly how the eternalist thinks of ‘now’: while claims that there is no particular time which is objectively now, utterances of the form ‘I’m hot now’ can perfectly well be true (and indeed translated into B-series expressions). They may go on to aver that, just as different spots in the road are no less real for being distant, events and objects at different times are equally real. Presentists disagree. They side with McTaggart on the B-series’ failure to account for the flow of time (they point out that ‘here’ doesn’t apparently “move locations”). They deny that the past and future is on the same ontological footing as the present: only the present is real. This has struck many philosophers as odd. As J. J. C. Smart put it, “Reality is not a property which anything can acquire. To be real is to be part of the universe” (1981, 142).

Though we seem to be able to put sense to the slogans of the presentist and eternalist, on reflection it is not always clear that their disagreement is genuine. Suppose we try making the foregoing slogan more precise: ‘For every x, x is present.’ This is equivalent to ‘it’s not the case that there exists something x such that x is not present’. But notice the tense of the verbs: if they are present tense, the claim is trivial. It amounts to the claim that everything that presently exists presently exists. If we read ‘exists’ instead as “disjunctively omnitemporal”, as in ‘it’s not the case that there existed, exists, or will exist something x such that x is not present’, the claim seems trivially false: plenty of things existed but no longer are present. Similar difficulties plague attempts to correctly formulate the eternalist position. Perhaps the presentist and eternalist are engaged in a merely verbal dispute.

One way of showing that the debate is in fact substantive may be to illustrate one or the other view’s ability to accommodate other philosophical problems (persistence, material constitution, responsibility, welfare, and so on). Identity, in particular, appears to be an important battleground for the eternalists and presentists.

**Identity**

*The Mission!*

Agent X has gone rogue and is attempting with her usual reckless zeal to expose the super-secret Identity Non-Proliferation Cabal (or “INPC”), an organization dedicated to spreading philosophical confusion. But she’s captured and subjected to The Process: her physical appearance, memory, and personality are altered so that she looks and behaves exactly like Joan Rivers — indeed, believes wholeheartedly that she is Joan Rivers. The INPC Directors find this hilarious. Just for kicks (since they hold Agent X’s Agency in such low esteem) they abduct the real Joan Rivers and subject her to the same procedure in the opposite direction, turning her into an Agent X look-alike/act-alike. They send Joan (as Agent X) back to the agency and Agent X (as Joan) back to Hollywood and bite their lips to keep from cracking up. Luckily, the Agency sees through the ruse and drags Agent X off the Red Carpet mid-celebrity-interview. Rogue agents must be punished. She protests: “Don’t you know who I am!? I’m Joan Rivers! I’m not an enemy of the state!” Someone in the van (who looks very much like Agent X), stares at her feet and feels guilty. . . .

Variations on this sort of story go back at least to Locke (1689/1975, Book II, Chapter 27), who imagined the consciousness of a prince and cobbler “switching bodies”. This showed, he thought, that the identity of a person consists in nothing but the memories and conscious experience, which are independent of the body. Many have

---

5 Here I follow Crisp’s (2003) formulation of the problem, also addressed in Lombard (1999) and this volume.

6 Such cases are discussed in much more detail in Williams (1970), Nozick (1981), and Parfit (1984), each of which is reprinted in Martin and Barresi (2003); Perry (1975) also collects many classic writings on personal identity.
objected to Locke’s argument. Joseph Butler, for example, pointed out that the argument is only compelling if we can be assured that the cobber is genuinely remembering the deeds of the prince, but Locke cannot assure us this unless he simply begs the question. Peter van Inwagen put essentially the same point in colorful terms (worth quoting at length):

I know a prince, Charlie, and a cobbler, Harry. One day the following seems to happen: Harry without warning begins to talk and act exactly as if he were Charlie. He exhibits both an unshakable conviction that he is Charlie and perfect knowledge of the most particular and intimate details of Charlie’s life. . . . (And, of course, all this happens, or seems to happen to Charlie, but the other way round: He loses his horsemanship but becomes able to mend shoes.) . . . Should I be forced by this strange occurrence to concede that a person is not after all a n organism and that the organism that had hitherto been “associated with” Charlie was now associated with Harry? Well, no, I certainly wouldn’t say that. Better to say nothing than to talk nonsense. And, really, in such a case there would be a great deal to be said for saying nothing. The only thing that would be clearly true would be that something wholly mysterious had happened. What should one say if the Eiffel Tower suddenly sprouted wings and flew away? Probably there is nothing one should say, beyond admitting that one had no idea what the explanation of this phenomenon is. (van Inwagen 1990, 187–188)

Countless further variants on these sorts of cases — not all of them so fanciful — have been discussed in the philosophical literature. They put pressure on different philosophical accounts of what it is to persist, how we can be responsible for our actions, and indeed, how we conceptualize ourselves.

Platitudes?

Something might seem fishy. If Lewis and company are correct that we never have a philosophical problem with identity, what (aside from the INPC) explains the persistence of puzzled philosophers? Perhaps our trouble devolves from the clashing of our characterization of identity with intuitions about how to apply the identity relation. We characterized identity above as the relation that everything bears to itself. This is reflexivity. Identity is also transitive: if Jones is Jekyll and Jekyll is Hyde, then Jones is Hyde. It is symmetric: if Jekyll is Hyde, then Hyde is Jekyll. Moreover, if Jekyll is Hyde, then whatever can be truly predicated of Jekyll can be truly predicated of Hyde.

Th is last principle — known variously as the Indiscernibility of Identicals and Leibniz’s Law — has been a longstanding source of puzzlement. A simple way of putting the same idea is that we can always substitute coreferring terms for each other without changing the truth-value of sentence (that is, salva veritate). Suppose the sentence ‘Jekyll is tall’ is true. If Jekyll is Hyde (if ‘Jekyll’ and ‘Hyde’ refer to the same person), then the sentence ‘Hyde is tall’ must also be true — the sentence is in effect saying the same thing or expressing the same proposition: that this guy (whatever we call him) is tall. But now consider the true sentence ‘Mary is in love with Jekyll’. When we substitute ‘Hyde’ for ‘Jekyll’, we seem to get a falsehood: Mary is not in love with Hyde — she fears him, in fact! What has gone wrong? Well, perhaps nothing. We can after all make some sense of the thought that Mary is in love with Hyde — for she is in love with Jekyll . . . and Jekyll is Hyde (her ignorance aside). She’s in love with that guy. Other examples are more recalcitrant. Take Quine’s chestnuts from

---

7 Though some reserve this name for the converse principle: that for all x and y, if x and y are indiscernible (i.e., if they have all the same properties), then they are in fact identical; this principle has historically been regarded as more controversial — so long as we take properties to be “sparse”. See Black (1952) and Hawthorne (2003, §2.2) for discussion.
8 Salmon (1981) offers an excellent discussion of cases of unknown identities. Examples like these go back at least to Frege (1969) and arguably to Eubulides’ paradox of the Masked Man: The masked man is your father; you don’t know who the masked man is; hence (paradoxically), you do not know who your father is (see Rescher 2001, 103).
“Reference and Modality”: Giorgione is Barbarelli and Giorgione was so-called because of his size; not so for Barbarelli. Nine is the number of planets; necessarily, nine is greater than seven; but the number of planets is not necessarily greater than seven. Quine’s approach to these apparent counterexamples to the Indiscernibility of Identicals was to identify certain semantic situations — typically those involving belief, modality, or where the names involved are not “purely referential” — in which the principle is fallaciously applied. He called these “opaque contexts” and much philosophical effort has been exerted trying to fully understand them.

Not surprisingly, the contrapositive of the Indiscernibility of Identicals (call it ‘the Non-Identity of Discernibles’) has also caused many headaches. It says that for any x and y, if x has any property y lacks, then x and y must be distinct. The same sorts of problems as we encountered above still apply: Jekyll seems to possess the property is loved by Mary while Hyde lacks it; so the principle wrongly implies that they are distinct. Similar gambits avail themselves: perhaps we’re again in an opaque context or no genuine property is expressed by ‘is loved by Mary’ — but then we’re saddled again with the problem of identifying opaque contexts or distinguishing the properties from the mere predicates.

The Problem of Change and Persistence over Time

The same reasoning behind the Non-Identity of Discernibles lurks behind the old and fertile dispute between Heraclitus and Parmenides. Parmenides argued that change was impossible. For if something changes by gaining or losing a property, it becomes “what it is not”. But that (Parmenides thought) is impossible. Reality is permanence; only appearances change. The slogans attributed to Heraclitus — that the world is change and that one cannot step twice in the same river — record prima facie disagreement.

But there is a sense in which Heraclitus agreed. We might thus accuse these two as engaging in a merely verbal dispute. Heraclitus’ dictum that the world is change can seem a bit ironic: the Heraclitan world is in fact one where no object changes. It’s a world of change only in the sense that there’s a constant stream of different objects that momentarily pop into and go out of existence, one after the other. You cannot step twice in the same river because as the water flows by, new rivers constantly come into and go out of existence. By way of analogy, consider a film: does it change? If we think of the film as a whole, it does in virtue of being composed of many different frames. If we focus instead on the individual frames, they remain static and separate.

Applying these images to ourselves initiates a worry. As Quine put it, “Undergoing change as I do, how can I be said to continue to be myself? Considering that a complete replacement of my material substance takes place every few years, how can I be said to continue to be I for the more than such a period at best” (1953, 65). Yesterday, I had the property of being sore from playing too much basketball. Today, I lack that property. Does the Non-Identity of Discernibles show that I am no longer the same person? How can I change and yet remain the same? The very idea looks like a contradiction!

Not so fast. We must distinguish between qualitative identity and what philosophers refer to as numerical identity. Imagine a pair of twins. Call them Rod and Todd. Rod and Todd are special in that not only are they genetically identical, they are physically identical in general down to the cell of skin and flake of dander (or anyway, we can suppose that they are at some point in their life). Rod and Todd are identical in the qualitative sense: they look exactly the same; they have all the same qualities (properties).

---

9 See Quine (1953, 140). Harold Noonan has suggested we identify a class of predicates "whose reference is affected by the subject term to which they are attached" “Abelardian” after the eleventh century philosopher Peter Abelard (1991, 188).

10 These reports (of unknown accuracy) come from Aristotel’s Physics (8. 3. 253b9) and Plato’s Cratylus (402a), respectively.
And yet they are different: while qualitatively identical, they are numerically dissimilar in just the sense that Rod is not the same person as Todd. The story of Rod and Todd’s adventures is a tale of two people, not one (hence “numerical” identity). We might imagine being friends with Rod and Todd and being put in situations where we weren’t sure whether we were still talking with the same person, even though we could be quite sure that whichever person this was, he had all of the same properties of Rod. Suppose, during halftime, Rod gets up to grab a beer. We might wonder of the person returning whether he is the same as Rod — whether Todd just happened to walk in a minute later with a beer. We can thus ask whether these “two” people — the person who got up and walked out and the person who walked back in (who are, for all intents and purposes, duplicates) — are really one and the same person. So numerical and qualitative sameness can come apart in at least one direction: qualitative identity does not imply numerical identity. But we still might wonder about the other? Does numerical identity require qualitative identity? Does the Non-Identity of Discernibles show us that nothing can change?

Most philosophers suppose not. Though as with the debates about time, they disagree as to how things persist through change. Ordinarily, we regard things as changing and yet not going out of existence all the time. My car recently got new tires. It didn’t at that point cease to be my car: some of its parts merely changed over. Intuitively, it seems that things can survive some change but not other (more significant) changes. Some philosophers — most famously, Descartes — attempt to solve the problem of personal identity by positing an immortal and unchanging soul that preserves our identity through change. The view faces several difficulties, not least of which is its inapplicability to “unsouled” objects.11 As Quine put it, even if we find it “agreeable to be driven . . . to believe in a changeless and therefore immortal soul as the vehicle of my persisting self-identity,” that solution won’t help us understand the persistence of purely-physical things like my car (ibid.).

We might attempt to repel the arguments against persistence through change by temporally-indexing the properties. I possess the properties sore-at-t1 and not-sore-at-t2 (as it were, eternally), but as these are different properties, there’s no problem.12 A related strategy invokes the space-time analogy mentioned above. We often ascribe incompatible properties to the same object by ascribing them to that object’s parts. I ask you if you’re warm enough. You reply that you’re warm and cold. Contradiction? No, you explain: while your trusty boots are keeping your feet warm, your hat is way too thin for these Idaho winters. Spatial parts can possess incompatible properties. Perhaps objects persist — or as defenders of this view like to say, “perdure” — in virtue of possessing different “temporal parts”. The perdurantist solves the apparent paradox of change by in much the way we solved the apparent paradox of warm and cold: when we predicate soreness to me, we predicate it to some of my previous temporal parts (and luckily withhold it from my present temporal parts).

So persistence through change is no problem. On this view, objects change simply by possessing different temporal parts (much like movies change by possessing different looking frames). They persist by being extended in both space and time.13 Some perdurantists thus call themselves “four-dimensionalists”.14 Though the

11 Perry offers a neat argument against recourse to the soul as the “vehicle” for our persistence in “The First Night” of his Dialogue on Personal Identity and Immortality: “If . . . identity of persons consisted in identity of immaterial unobservable souls, then judgments of personal identity of the sort we make every day whenever we greet a friend or avoid a pest are really judgments about souls . . . . But if such judgments were really about souls, they would all be groundless and without foundation. For we have no direct method of observing sameness of soul . . . . But our judgments about persons are not all simply groundless and silly, so we must not be judging of immaterial souls after all” (1978, 11–12).

12 Well, no problem with Non-Identity of Discernibles; some worry about the problem of “temporary intrinsics” (see Lewis 1988).

13 These spatiotemporally-extended objects are sometimes called “space-time worms”, though I much prefer the image of the “space-time salami”.

6
perdurantist has a natural ally in the eternalist, the views are separable. Spacetime worms could exist on the Growing Block model of time and perhaps even on a presentist view of time.

Many find the perdurantist solution philosophically expensive and unwarranted. It implies, for example that I’m am not now “wholly present”. That seems odd. Th ings just endure change — they gain and lose parts as time flows by, but that does not make them “spread out through time” in some special way. Th ese “endurantists” can reply to the problem of persistence through change by taking a page from the presentist’s playbook: so long as we “have a care for tense”, the Non-Identity of Discernibles does not generate a contradiction. While I had the property of soreness, I am now not sore. Th ese philosophers deny that change poses any real problem for persistence. Thus, like the perdurantists, though endurantists have a natural ally in the presentists, most regard their views as independent. And as with the debate between the presentists and eternalists, the debate between the endurantists and perdurantists (or the “3Ders” and “4Ders”) has been supposed to be merely verbal. Each camp, of course, opposes this conclusion and supposes their view to deal more effectively with various metaphysical paradoxes.\(^{16}\)

**The Self**

How do questions about the metaphysic of persistence bear on basic questions about what sorts of things we are? How do different approaches to the latter question bear on the former? Thi is is a rich area of philosophical inquiry. We can dip into it by considering how another one of our “platitudes” about identity — this time, transitivity — buts up against our intuitions.

Take the Ship of Theseus.\(^{17}\) While at sea, its ingenious crew carries out repairs, discarding worn wooden planks and replacing them with aluminum. Equally ingenious mariners somehow salvage these cast-off parts — every last one — and eventually rebuild the wooden ship. So we have a wooden ship which has the very same parts as Theseus’ original ship and an aluminum ship which seems to have survived each change to the original. Now, it sure seems plausible that either of the resulting ships are the original ship (after all, if the wooden ship did not exist, the aluminum ship would clearly be regarded as the same ship; ditto for the wooden ship if the aluminum ship did not exist). But they cannot both be identical to original ship. For suppose they were: if the Original is the Wooden and Original is the Aluminum, then the Wooden is the Aluminum. But that’s not so: whatever they are identical to, they are not identical to each other — they are, after all, two ships. So which is the original ship?

There’s much to say about this case — but what does it have to do with the self? Well, for one, just as we can imagine a ship “fissioning” into two things, we can imagine people fissioning. Perhaps the INPC transplants the two hemispheres of your brain to two brainless bodies or perhaps, to use Parfit’s famous example (1984), the matter-transporter breaks down and creates two whole duplicate copies of you. Either way, you cannot be identical two both of them for the same reason that the original ship cannot be both the wooden ship and the

---

\(^{14}\) See Sider (1997), (2001) and Hawley (2001) for extended explications and defenses of this family of views. Its popularity notwithstanding, some find the term ‘Four-Dimensionalism’ to be non-ideal; for it’s not yet clear that space has only three dimensions (or time only one), or even that either has a definite number.

\(^{15}\) See, for example, Wiggins (1980), Lombard (1986), and Van Inwagen (1990) on the desirability of things persisting “wholly present”.

\(^{16}\) Sider (2001,) offers a vigorous rebuttal of the “metaphysical skeptic” and illustration of four-dimensionalism’s virtues.

\(^{17}\) The problem is mentioned first in Plutarch (Life) and returned to modern philosophical consciousness by Hobbes (De Corpore); I loosely follow Chisholm’s (1976, Chapter 3) presentation.
aluminum ship: two people cannot be one. But if the procedure is symmetric, it seems intolerably arbitrary to claim that only one or the other person is you. Nor does it look very plausible to say that neither is you.\textsuperscript{18}

Parfit offered the radical suggestion that identity is not “what we care about” when it comes to our survival. Because the identity relation is transitive, I cannot be identical to the duplicates; but I can share their memories, beliefs, desires, personalities, goals, projects, and so on. From my point of view, things are just as good as being identical (indeed, I might not even know about my “competition”). When I consider what I care about in survival, it is my psychology.

Parfit’s suggestion has generated an immense literature.\textsuperscript{19} It is itself part of an immense literature exploring the idea that the criteria for personal identity over time must be (at least partly) psychological.\textsuperscript{20} Specifically what psychological features matter or how these features are realized are deep and interesting questions. As we mentioned above, Descartes believed that he was a constant immaterial soul. Hume claimed to find no such evidence of this: all he could recognize were a bundle of different perceptions.

Descartes also nicely illustrates a recent concern over the psychological approach. He thought, recall, that the mind and body were distinct substances. He could doubt the existence of the latter but not the former, so by the Non-Identity of Discernibles (pace concerns already raised), they must be distinct. But then what is my relation to the organism sitting in my chair? Th at creature thinks, doesn’t it? Are two things thus thinking these same thoughts? Th at is strange. Eric Olson (1997; 2002; 2003) has been leading a campaign to rid the philosophy of personal identity of psychology. He claims that we are thinking human animals — that we were once fetuses and may someday end up in a persistent vegetative state. In many ways, this debate is orthogonal to the debate between the perdurantist and the endurantist, but how we approach one debate will likely influence our approach to the other.

As with the issues of time and identity, our philosophical puzzlement about the nature of the self is itself puzzling. We ask what is the self? but it might seem that we can have no better access to anything but ourselves. For those of us who aren’t quite sure what we are, how can we not be? Do we misunderstand the question? Are we diving into a shallow puddle thinking it a deep pool?

**RECENT WORK ON TIME AND IDENTITY: THE ESSAYS**

I’ve claimed that the concepts of time and identity are intimately connected. Nevertheless, the essays are divided (rather roughly) into four sections corresponding to their respective centers of gravity: *Time, Identity, The Self,* and *Death.*\textsuperscript{21} Here’s a brief tour.

\textsuperscript{18} Suppose there’s a lag in the duplication: the transporter “moves” me down to the planet’s surface and then five minutes later malfunctions and produces a duplicate. Do I then go out of existence? Do I suffer “death from competition” (as Perry 1978 put it)?

\textsuperscript{19} Many oppose the suggestion that survival does not involve identity. As Lewis puts it, “Suppose I wonder whether I will survive the coming battle, brainwashing, brain transplant, journey by matter-transmitter, purported reincarnation or resurrection, fission into twins, fusion with someone else, or what not. What do I really care about? . . . What matters in survival is survival. If I wonder whether I will survive, what I mostly care about is quite simple. When it’s all over, will I myself — the very same person now thinking these thoughts and writing these words — still exist? Will any one of those who do exist afterward be me? In other words, what matters in survival is identity — identity between the I who exists now and the surviving I who will, I hope, still exist then” (Lewis 1976, 55). Lewis and Sider both solve the “fission problem” by employing temporal parts.

\textsuperscript{20} Indeed, the possibility of fission was first conceived (as early as Butler) as an objection to the psychological approach to personal identity. We might think of Parfit’s particular radical spin on the psychological approach to survival as a way of disarming this objection.

\textsuperscript{21} Dividing the essays in the present volume into sections apparently disregards this fact! Organization often has the effect of imposing more order than exists — best, then, to view these containers as porous.
The essays in this section address the metaphysics of time and the conceptual links between time and freedom. The first two—by Lynne Rudder Baker and Lawrence Lombard—express different kinds of dissatisfaction with the debate between the eternalist and the presentist.

Baker, in “Temporal Reality” (Chapter 2), argues that each alone fails to cohere with either physics or human experience. The A-theory, while in accord with our experience of time, appears incomplete or incompatible with modern physics. The B-theory, though apparently required for physics, cannot make sense of either the inexorable “flow” of time or the fact that you have less than a year to live or that the earth is now billions of years old. Baker proposes an intriguing theory of time—the BA theory—which takes the B-series as basic. She writes, “In the absence of self-conscious beings, events occur (tenselessly) at various times, and some events are (tenselessly) later than others. But there is no ongoing now” (13). Her account makes the A-series facts relative to the experiences of self-conscious beings: without such beings, “there are no A-series” (14). But rather than taking this feature as implying that the A-series is unreal or “merely mind-dependent”, she argues that the existence of self-conscious beings is a genuine feature of reality that has implications for other general features of reality. She closes her essay with an extended discussion of the implications of the BA-theory for the relation between time and existence.

This relation looms large for Lombard, in “Time for a Change: A Polemic Against the Presentism/Eternalism Debate” (Chapter 3). He assimilates the ancient debate about change between Heraclitus and Parmenides to the contemporary debate between the presentists and the eternalists: perhaps they too are engaged in a merely verbal dispute. As we have seen above, it is surprisingly difficult to explicate these views in a way that generates a substantive dispute. Lombard interrogates glosses by Merricks (1995), Zimmerman (1998), and Sider (2001), concluding that they all founder on an equivocation of tense. “If all the relevant verbs [in the definitions of eternalism and presentism] are in the present tense, there is no substantive dispute. . . .” (51).

Another way that eternalists and presentists might attempt to spell out their differences involves the “fixity” of the future. The eternalist’s claim that past, present, and future entities are all equally “real” might usefully parlay into the claim that those facts are fixed—a claim the presentist may deny, at least about the future. Whether or not this suggestion will satisfy “skeptics” like Lombard, it raises an interesting question about the relation of our metaphysics of time and our view of our freedom. Suppose the future is fixed; then if I sprain my ankle on the basketball court tomorrow, it is true now that that event will come to pass and there is nothing I can do to prevent it. So the argument for fatalism goes. A similar issue arises in discussions of time travel. Kurt Gödel provided solutions to Einstein’s field equations that vindicated the possibility of “closed timelike curves”, elevating time travel from entertaining fiction to tantalizing possibility. Many philosophers were skeptical: they worried about logical paradoxes that might follow from meddling time travelers bent on bringing about their own non-

---

22 Baker suggests that this is true of the Growing Block theory as well; see Hinchliff (1996) and (1998) for a defense of the compatibility of presentism with relativity.

23 The possibility of time travel raises all manner of fascinating problems (and entertaining stories!) which bear on issues of time and identity. Does one have to be an eternalist to believe in time travel (see Keller and Nelson 2001; Sider 2001, §7.2)? What do we say about the time traveler conversing with herself: how many people are talking? Might there be situations where a time traveler is only indeterminately identical with someone who steps into a time machine (Slater 2005)?

24 Palle Yourgrau has written extensively and illuminatingly on Gödel’s views; see, for example, his (1999); compare Earman (1995).
existence. David Lewis set many minds at ease in his seminal (1976) essay: time travelers can’t, say, kill their grandfathers for the commonplace reason that they didn’t. But not everyone got on board. Kadri Vihvelin (1996) argued that this treatment downplayed the strange inability of time travelers to do what we otherwise think them perfectly capable of doing. In his “Context, Conditionals, Fatalism, Time Travel and Freedom” (Chapter 4), John Carroll offers a contextualist account of counterfactual conditionals designed to sort out the dispute between Vihvelin and Sider (2002) on the abilities of time travelers and diagnose the fatalist’s argument. On this account, sentences like ‘Tim the Time Traveler cannot kill his Grandfather’ are true, but in a weak and non-puzzling sense.

Other worries beset the presentist if we take him at his word that no past and future individuals exist. How do we yet claim that Lincoln was shot by Booth and that the Earth will orbit the sun? Mark Hinchliff addresses this question in his “The Identity of the Past” (Chapter 5), locating “Property Presentism” as the principle responsible for these worries. Roughly speaking, this principle says that if something has a property, that something exists. Put that way, it seems impossible to deny. Hinchliff weakens its hold over presentists by exploring the analogous relation between actualism (the view that only actual entities exist) and property actualism. He argues that the analogy breaks down when we take a certain view about the difference between the past and the future (on some presentist’s lights): while the future, like the merely possible, is “irreducibly general” (resisting cross-time or cross-world identifications), the past “has a full identificatory structure of particular instances underlying it” (12). This point opens the door to an extensive batch of apparent counterexamples to property presentism. And as its fortunes wane, so do the force of the problems surrounding referring to past people and events.

Identity

Identity may be simple, but as we’ve seen, there’s plenty of room for argument over how to apply that concept to problems in metaphysics and the philosophy of language. We begin with the latter. Recall that one of the “platitudes” regarding identity was the principle of substitutivity or the indiscernibility of identicals. While substitutivity of identicals fails in opaque contexts, it is generally thought that names are “purely referential” — for example, when the names are “rigid designators” or “directly referential” — the substitutivity principle is true. But things may not be as simple. In “Identity through Change and Substitutivity Salva Veritate” (Chapter 6), Reinaldo Elugardo and Robert J. Stainton present a new puzzle about substituting corefering names into sentences describing accidental change. The puzzle arises from the fact that an object’s name often changes along with its other properties. Toronto, for example, once had small, waterlogged streets and was thus called ‘Muddy York’. But its streets were drained and its name changed to ‘Toronto’. We might say that ‘Muddy York evolved into Toronto’. But if we grant the substitutivity principle and grant that the names ‘Muddy York’ and ‘Toronto’ refer to the same thing, “we arrive at a contradiction: It’s not the case that Toronto evolved into Toronto!”

Their solution to this paradox abjures construing “_____ evolved into _____” as an opaque context as philosophically expensive and unwarranted. Instead, Elugardo and Stainton diagnose the problem as devolving from the polysemy of the names involved: we use names sometimes to refer to an object over time (a continuant, for example) and sometimes to an object at a time. The paradox thus devolves from a subtle shift in lexical role of ‘Toronto’ and ‘Muddy York’. Though their solution professes metaphysical neutrality, it brings out a point of

---

25 In the sense of Kripke (1980).
methodological significance for metaphysics: the possibility of some kind of reference-shifting need not generate opaque contexts. The relation ‘___ evolved into ___’ “just does apply to the object however conceived” (25).

The remaining essays in the volume address personal identity from a number of different angles. At a metaphilosophical level, Ned Markosian — in “Identifying the Problem of Personal Identity” (Chapter 7) — suggests that the standard way of putting the problem improperly biases the ensuing debate in favor of four-dimensionalism. Philosophers will often put the problem this way: under what conditions are person x at time t₁ and person y at time t₂ in fact the same person? But to what entities do the phrases ‘person x at t₁’ and ‘person y at t₂’ refer? Four-dimensionalists (perdurantists) have a ready answer: to the temporal parts of person x and person y. Three-dimensionalists (endurantists) have no truck with this — for they deny that persons have temporal parts. They prefer to ask after the conditions under which something that is a person at t₁ is the same person as something that is a person at t₂. But as we have seen, invoking sortal identity in order to state the problem incurs a heavy philosophical burden.

Drawing upon a new theory of property instantiations, Markosian proposes a new way of putting the problem that he claims levels the playing field for the 3Der and 4Der. Properties, he notes, are often instantiated for extended periods of time: the sun has been hot for several billion years; leaves stay green in summer and turn red in the fall; the number seven has always been prime. Call these instantiations “episodes”. Now we can ask some questions about whether this episode is the same as that. In particular, we can ask of persons, What are the conditions under which an instance of personhood at t₁ is part of the same episode of personhood as an instance of personhood at t₂? Th is characterization help make sense of some of the vexing problems facing the 3Ders (the Fission Problem, the Time Travel Problem, and the Fetus/Corpse Problems among others). Though particularly welcome for the 3Der, success here should be regarded as good news all around, as it seems preferable to not presuppose one way of thinking about a problem by merely stating it (a theme Markosian sounds more than once in his paper).

In a similar spirit, Neal Tognazzini in his “Persistence and Responsibility” (Chapter 8) rebuts several arguments for the incompatibility of claim that we are responsible agents with perdurantism. As noted above, ascriptions of moral responsibility often seem to depend on ascriptions of identity among agents. Some philosophers allege that the metaphysics of perdurance leaves no room for the existence of responsible agents. Consider our Agent X again, after her imprisonment but before she is subjected to the (Joan Rivers) Process. To be held responsible for her actions, it clearly must be the case that the Prisoner is Agent X. But critics of perdurantism point out that the object in custody is not all of Agent X. Her captured temporal part differs from her thieving temporal part: they are numerically distinct. But if distinct, how can we hold the captured person stages responsible for crimes they literally didn’t commit? Tognazzini replies that this and other objections simply represent prejudice against the perdurantist way of conceiving of numerical identity. He writes: “The Perdurantist can quite plausibly claim that what is required for an attribution of moral responsibility to be appropriate is not that ‘the self-same entity’ be ‘wholly present’ at both times, but rather that the same person be present (but not wholly) at both times” (6–7). And the perdurantist can easily make sense of this fact. In responding to this and similar objections, Tognazzini fleshes out perdurantism’s ability to integrate ascriptions of moral properties to continuants.

Geoffrey Gorham, on the other hand, seeks to pin down one famous philosopher’s metaphysical commitments. In “Descartes on Persistence and Temporal Parts” (Chapter 9), Gorham contends that the best
way to reconstruct Descartes’ argument for the immortality of the soul makes use of the modern resource of perdurantism. His interpretation begins with a suggestive passage from the Third Meditation:

For a lifespan can be divided into countless parts each completely independent of the other, so that it does not follow from the fact that I existed a little while ago that I must exist now, unless there is some cause which as it were creates me afresh at this moment — that is, preserves me.

Th is apparent independence of different “momentary” souls strongly suggests a perdurantist reading of persistence: perhaps we should conceive of a Cartesian soul as composed of countless temporal parts, each dependant on some external causal influence for their existence. However, this interpretation raises some very tricky problems of interpretation for Descartes. One might think, for example (following Bennett 2001), that a genuine substance cannot possess distinct substantial parts; that it would be a “mere pseudo-substance” (vol. 1, 98). Gorham finds precedent in Descartes for denying this doctrine. More worrisome, however, is Descartes’ famous insistence upon the simplicity of the soul — as we’ve seen, this simple, unchanging soul is one way of attempting to secure an agent’s numerical identity in the face of qualitative change. The solution to this problem, Gorham suggests, lies in the unchanging individual essence behind each thinking substance. Th is is the Cartesian ego that persists unchanged — and indeed may be argued to exist out of time and thus, in a sense, “immortal by its very nature” (12).

**The Self**

The essays in this section, though they could easily have been lumped together with those of the previous, share a focus on our first-person experience — our understanding of our selves. What sort of things are we? Are we immortal souls like Descartes thought? Are we instead bundles of perceptions and thoughts? Do we even need to settle these issues to enjoy a conception of ourselves?

In “Persons, Animals and Human Beings” (Chapter 10), Harold Noonan leans on an intriguing view of first-person reference to further articulate his approach to personal identity and personhood.26 As we’ve seen, we might ask very generally, “What changes can a person can survive? What changes will terminate a person’s existence?” Noonan argues that the indexical formulation of the problem is more basic: “Our interest in personal identity is fundamentally an interest in our own identity.” On this view, persons are just the objects of first-person reference. Armed with this simple conception of persons, Noonan suggests that defenders of the Psychological Approach to personal identity can rebut Olson’s “Th inking Animal” problem (what he calls “Too Many Minds Objection”). Perhaps we should admit that each of us “is” an animal “in the sense of coinciding with one and being constituted of the same matter as one — but this ‘is’ is the ‘is’ of constitution, not identity” (15).27 But even if persons and human animals coincide in this manner and both think “I-thoughts”, it does not follow that their thoughts are about different thinkers. Noonan thus dissolves the skeptical difficulties associated with the Too Many Minds Objection — “Both the person and the animal can know that their utterance of ‘I am a person’ is true” (ibid.) — for ‘I’-thoughts always refer to the person thinking them. His essay concludes with an extended defense of this approach from Olson’s (2002) objections.

Jenann Ismael also takes up the issue of first-person reference in “Me, Again” (Chapter 11). She begins with Anscombe’s objection Descartes’ argument that he is an immaterial thinking substance — roughly, that

---

26 For previous statements of this view, see Noonan (1998) and (2003).

27 Th is would seem to represent a shift from Noonan’s earlier views on constitution as identity (see his 1993).
Descartes cannot guarantee that the same referent is picked out in different ‘I’-thoughts without presupposing some criterion of identity for selves (which is what’s under discussion). But Ismael points out that unlike other indexicals (like ‘here’ and ‘now’) which can be mistakenly substituted for one another, “it’s hard to make sense of the idea of mistakenly intersubstituting someone else’s I-occurrence for one of our own” (7). Criteria of identity “are not employed in reidentification” (ibid.). We may thus manage to refer to ourselves (both at a time and over time) without possessing an explicit concept of the Self. Th is realization has suggested to some philosophers that the self must in fact be a primitive. Ismael argues that we can avoid this conclusion by seeing the concept of the self as constituted by these reidentifications (rather than being presupposed by them).

John Perry — in “Selves and Self-Concepts” (Chapter 12) — likewise resists the thought that the self is something mysterious. He proposes a “Straightforward Theory” of the self, offering the following analogy: a neighbor is just a person thought of under the relation of living next door to someone; likewise, a self is just a person thought of under the relation of identity. “Self is to identity, as neighbor is to living next door to” (1). On this view, philosophical perplexities about the self stem not from the self being a special kind of object (selves, like neighbors, are just persons), but by the unique epistemic structure of this concept. Normally, we integrate knowledge about ourselves gained from external sources and from “internal” sources (what Perry calls “normally self-informative ways of knowing about a person”) which are usually immune to error. Th e first ‘normally’ stems from cases like the amnesiac war-hero who learns about himself from written accounts without knowing that those written accounts describe his actions; the second ‘normally’ owes to the recognition that our modes of gaining information about our surroundings and such often presuppose contingent facts about how our senses are “hooked up.” Just how we integrate information from these sources may lurk behind the different public identities we assume (teacher, student, football fan, and so forth). Our identity consists in those parts of our self-concept that are central — those things we cannot imagine ourselves not being. Th is suggests to Perry a sort of “Bundle of Bundles Theory” of the self on which different “competing centers of agency” jockey for position. Such an account, he argues, makes good sense of our not-always-coherent mental life.

A natural way of elaborating Perry’s “centers of agency” is the desire satisfaction account of well-being (pursued by Parfit). We often define ourselves by our desires or long-term plans; we look forward to having those desires fulfilled — that benefits us. But our desires often change: what I wanted ten years ago is not what I want now. In “Ex Ante Desire and Post Hoc Satisfaction” (Chapter 13), H. E. Baber addresses this problem for the desire satisfaction model of well-being, claiming that the satisfaction even of preferences I can no longer identify with does benefit us, even if we are not better off, even if we don’t appreciate our desires being satisfied. Th is stance leaves us better able to make sense of the obscure artist who only benefits after her death. Baber defends the view that we benefit when the desires are satisfied.

28 Bertrand Russell expressed a similar sentiment when he quipped “I think, therefore I am’ says rather more than is strictly certain. It might seem as though we were quite sure of being the same person today as we were yesterday, and this is no doubt true in some sense. But the real Self is as hard to arrive at as the real table and does not seem to have that absolute, convincing certainty that belongs to particular experiences. When I look at my table and see a certain brown colour, what is quite certain at once is not ‘I am seeing a brown colour’, but rather, ‘a brown colour is being seen’” (Russell 1912, 19).
29 Perry’s use of ‘normally’ here recalls Daniel Dennett’s classic paper “Where Am I?” (chapter 17 of Dennett 1981) in which he describes having his brain transplanted to a vat where it controls his brainless body by remote. We might imagine performing this procedure on two patients and putting the feeds on a switch. Perhaps it would make sense for one patient to express doubt that she was experiencing her own sensations. Or compare the film “Being John Malkovich”.

13
Death

Just as we want to speak of posthumous benefits, we may want to account for posthumous harms. Rather than achieving renown after death, our obscure artist’s work may be stolen or maligned. Perhaps death itself is a harm. Defenders of this position face an ancient challenge: how could there be a harm without a subject? As Epicurus famously wrote in his Letter to Menoeceus: "so long as we exist, death is not with us; but when it comes, then we do not exist." Many have found this paradoxical: we clearly want to say that death is an evil, but can we coherently claim that we are harmed even when we do not exist?

 Construing death as an evil seems to place some constraints on our metaphysics: after my death, the presentist will claim that I do not exist to be benefited or harmed. Ben Bradley claims in his “Eternalism and Death’s Badness” (Chapter 14) that we must be eternalists to make sense of the cross-temporal relation between a person and that person’s death. While they do not exist simultaneously, they do both exist. He spends the bulk of his essay responding to arguments offered by Harry Silverstein (1980) and (2000) that death is an “atemporal” harm to the person who has died. Bradley contends in contrast (and in a spirit similar to Baber) that death is bad for a person when that person would have been living a good life had death not occurred.

 Silverstein finds this perplexing. He addresses Bradley’s criticism in his essay “The Time of the Evil of Death” (Chapter 15), sharpening and expanding his previous view that the question of when is S’s death an evil for S? is exhaustively answered by stating when S was born and when S died. There is nothing further to say. He points out that Bradley’s view seems to entail some strange conclusions: for example, that we ought to celebrate the point at which a dead loved one would have died had he or she not died the death he or she in fact died — as we might celebrate a loved one’s release from prison.

 Barbara Levenbook too wishes to accommodate the intuition that death is and evil (ceteris paribus). And she makes clear what is at stake: the very intelligibility of our possessing a “right to life” plausibly depends on our ability to explain how death counts as a frustration of interests. Addressing the Epicurean challenge in the guise of what she calls “The Retroactivity Problem” (Chapter 16) and also developing previous work (1984), Levenbook articulates a series of principles which give rise to the problem. Ultimately, the issue depends more on moral rather than metaphysical principles — in particular, a substantive theory of the “good for” and “bad for” relations. Thus, it would seem out intuitions about death carry with them more than merely metaphysical implications.

REFERENCES

Chisholm, Roderick (1976) Person and Object. La Salle: Open Court.


Rescher, Nicholas (2001) Paradoxes: Their Roots, Range, and Resolution. La Salle: Open Court.


