

# Philosophical Papers

---

Volume I

DAVID LEWIS

New York Oxford  
OXFORD UNIVERSITY PRESS  
1983

Copyright © 1983 by David Lewis

Library of Congress Cataloging in Publication Data

Lewis, David, 1941-  
Philosophical papers.

Includes index.

1. Philosophy—Addresses, essays, lectures.

I. Title.

B29.L49 1982 110 82-22551

ISBN 0-19-503203-9 (v. 1)

ISBN 0-19-503204-7 (pbk. : v. 1)

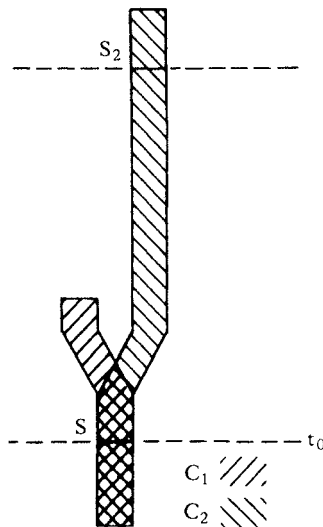
Printing (last digit): 9 8 7 6 5 4 3

Printed in the United States of America

*Postscripts to*  
 “Survival and Identity”

A. TWO MINDS WITH BUT A SINGLE THOUGHT

Derek Parfit rejects my attempt to square his views (which are mine as well) with common sense.<sup>1</sup> He objects that before I bring off the reconciliation, I must first misrepresent our commonsensical desire to survive. Consider a fission case as shown. I say there are two continuant persons all along, sharing their initial segments. One of them,  $C_1$ , dies soon after the fission. The other,  $C_2$ , lives on for many years. Let  $S$  be a shared stage at time  $t_0$ , before the fission but after it is known that fission will occur. The thought to be found in  $S$  is a desire for survival, of the most commonsensical and unphilosophical kind possible. Since  $S$  is a shared stage, this desire is a shared desire. Certainly  $C_2$  has the survival he desired, and likewise has what we think matters: mental continuity and connectedness (the R-relation) between  $S$  and much later stages such as  $S_2$ . But how about  $C_1$ ?



<sup>1</sup>“Lewis, Perry and What Matters,” in Amélie Rorty, *The Identities of Persons* (Berkeley: University of California Press, 1976), pp. 91–96.

I wrote that "if common sense is right that what matters in survival is identity . . . , then you have what matters in survival if and only if your present stage is I-related to future stages" where stages are I-related iff they belong to some single continuant person (page 59). If that is right, then  $C_1$  has what he commonsensically desired. For  $C_1$ 's stage  $S$  at time  $t_0$  is indeed I-related to stages far in the future such as  $S_2$ . These stages are I-related via the person  $C_2$ .—"But isn't this the *wrong* person?" says Parfit.  $C_1$  himself survives only a short time. The one who lives longer is another person, one with whom  $C_1$  once shared stages. If his desire is satisfied by this vicarious survival, it cannot really have been a commonsensical desire to survive.

If  $C_1$  really had the commonsensical desire that he himself—the continuant person  $C_1$ —survive well into the future, then I grant that his desire is not satisfied. But I don't think he could have had exactly that desire. I said that the desire found in  $S$  was to be *of the most commonsensical and unphilosophical kind possible*. And there is a limit to how commonsensical one's desires can possibly be under the peculiar circumstance of stage-sharing.

The shared stage  $S$  does the thinking for both of the continuants to which it belongs. Any thought it has must be shared. It cannot desire one thing on behalf of  $C_1$  and another thing on behalf of  $C_2$ . If it has an urgent, self-interested desire for survival on the part of  $C_1$ , that very thought must also be an urgent, self-interested (and not merely benevolent) desire for survival on the part of  $C_2$ . It is not possible that one thought should be both. So it is not possible for  $S$  to have such a desire on behalf of  $C_1$ . So it is not possible for  $C_1$  at  $t_0$  to have the straightforward commonsensical desire that he himself survive.

If  $C_1$  and  $C_2$  share the most commonsensical kind of desire to survive that is available to them under the circumstances, it must be a plural desire: *let us survive*. Now we must distinguish two different plural desires: existential and universal, weak and strong.

(weak) Let at least one of us survive.

(strong) Let all of us survive.

Because these desires are plural instead of singular, they are not perfectly commonsensical. Because they are put in terms of survival of continuants rather than relations of stages, they are more commonsensical than the "philosophical" desire for R-relatedness of one's present stage to future stages.

If  $C_1$ 's (imperfectly) commonsensical desire for survival is predominantly the weak desire, then my reconciliation goes through. For  $C_1$ 's weak desire is satisfied even though it is his stage-sharer rather than himself who survives. The weak desire is indeed equivalent to a desire for I-relatedness to future stages. Then if I am right that the I-relation is the R-relation, it is equivalent also to the desire for R-relatedness to future stages.

If  $C_1$ 's desire is predominantly the strong desire, on the other hand, it is not satisfied. Then his desire for survival is not equivalent to the "philosophical" desire

for R-relatedness to future stages, and my reconciliation fails. (However, the strong desire is equivalent to a more complicated desire concerning R-relatedness of stages.) But should we say that  $C_1$  has the strong desire, and that since it is not satisfied, he does not have what commonsensically matters in survival? I think not. For if we say that of  $C_1$ , we must say it also of  $C_2$ . If one has the strong desire, both do. The strong desire is no more satisfied for  $C_2$  than it is for  $C_1$ . But it seems clear that  $C_2$ , at least, *does* have what commonsensically matters in survival.

It is instructive to consider a system of survival insurance described by Justin Leiber, in *Beyond Rejection*.<sup>2</sup> (But let us imagine it without the risks and unpleasantness that Leiber supposes.) From time to time your mind is recorded; should a fatal accident befall you, the latest recording is played back into the blank brain of a fresh body. This system satisfies the weak desire for survival, but not the strong desire. Let  $S$  at  $t_0$  be the stage that desires survival and therefore decides to have a recording made; the fission occurs at the time of recording;  $C_1$  dies in an accident not long after;  $C_2$  survives. The only extra peculiarities, compared with a simple case of fission, are that  $C_2$  is interrupted in time and undergoes a body transplant. If this system would fairly well satisfy your desire for survival—or if your misgivings about it concern the body transplant rather than the fission—then your desire is predominantly the weak desire.

So far, I have supposed that  $C_1$  and  $C_2$  at  $t_0$  already anticipate their fission. Now suppose not. Now, cannot they share the perfectly commonsensical singular desire: let *me* survive? After all, the desire to be found in the stage  $S$  in this case is no different from the desire that would be there if  $S$  were what it takes itself to be: a stage of a single person with no fission in his future. I agree that  $C_1$  and  $C_2$  have the singular desire. But it is not a desire that can be satisfied, for it rests on the false presupposition that they are a single person. The “me” in their shared thought (unless it refers to the thinking stage) has the status of an improper description. It cannot refer to  $C_1$  in  $C_1$ 's thought and to  $C_2$  in  $C_2$ 's thought, for these thoughts are one and the same. But their desire to survive *is* satisfied; at least  $C_2$ 's is, and  $C_1$ 's is no different. Therefore their desire for survival cannot consist only of their unsatisfiable singular desire. They must have the weak plural desire as well, despite the fact that they don't anticipate fission. And so must we. Doubtless we seldom have it as an occurrent desire. But many of our urgent desires are not occurrent, for instance your present desire not to suffer a certain torture too fiendish for you to imagine.

(At this point the reader of “Attitudes *De Dicto* and *De Se*” (in this volume) may wonder how well I have learned my own lesson. There I taught that desire is a relation of wanting-to-have—take this as indivisible—that the subject bears to a property. Why can't  $C_1$  and  $C_2$  bear the very same wanting-to-have relation to the very same property of surviving, so that they think the very same thought, and yet each thereby desire his own survival? But recall that the subject that wants-to-have

<sup>2</sup>(New York: Ballantine, 1980).

properties was taken to be a stage, not a continuant. (See pages 143, 146.) Under this analysis, my point is that S's wanting-to-have the property

being such that the unique continuant of which it is a stage survives

is an unsatisfiable desire. That is so whether we think of it as a desire of S's or, more naturally, as a desire of  $C_1$  and  $C_2$ . S had better want survival on behalf of  $C_1$  and  $C_2$  by wanting to have a different property:

being such that some continuant of which it is a stage survives.

This is the satisfied desire for survival that  $C_1$  and  $C_2$  share.)

### B. IN DEFENSE OF STAGES<sup>3</sup>

Some would protest that they do not know what I mean by "more or less momentary person-stages, or time-slices of continuant persons, or persons-at-times". Others do know what I mean, but don't believe there are any such things.

The first objection is easy to answer, especially in the case where the stages are less momentary rather than more. Let me consider that case only; though I think that instantaneous stages also are unproblematic, I do not really need them. A person-stage is a physical object, just as a person is. (If persons had a ghostly part as well, so would person-stages.) It does many of the same things that a person does: it talks and walks and thinks, it has beliefs and desires, it has a size and shape and location. It even has a temporal duration. But only a brief one, for it does not last long. (We can pass over the question how long it can last before it is a segment rather than a stage, for that question raises no objection of principle.) It begins to exist abruptly, and it abruptly ceases to exist soon after. Hence a stage cannot do everything that a person can do, for it cannot do those things that a person does over a longish interval.

That is what I mean by a person-stage. Now to argue for my claim that they exist, and that they are related to persons as part to whole. I do not suppose the doubters will accept my premises, but it will be instructive to find out which they choose to deny.

First: it is possible that a person-stage might exist. Suppose it to appear out of thin air, then vanish again. Never mind whether it is a stage *of* any person (though in fact I think it is). My point is that it is the right sort of thing.

Second: it is possible that two person-stages might exist in succession, one right after the other but without overlap. Further, the qualities and location of the second at its appearance might exactly match those of the first at its disappearance. Here

<sup>3</sup>On this topic I am much indebted to discussions with Saul Kripke and with Denis Robinson. Kripke's views on related matters were presented in his lectures on "Identity through Time," given at Princeton in 1978 (and elsewhere); Robinson's in "Re-Identifying Matter," *Philosophical Review* (forthcoming).

I rely on a *patchwork principle* for possibility: if it is possible that X happen intrinsically in a spatiotemporal region, and if it is likewise possible that Y happen in a region, then also it is possible that both X and Y happen in two distinct but adjacent regions. There are no necessary incompatibilities between distinct existences. Anything can follow anything.

Third: extending the previous point, it is possible that there might be a world of stages that is exactly like our own world in its point-by-point distribution of intrinsic local qualities over space and time.

Fourth: further, such a world of stages might also be exactly like our own in its causal relations between local matters of particular fact. For nothing but the distribution of local qualities constrains the pattern of causal relations. (It would be simpler to say that the causal relations supervene on the distribution of local qualities, but I am not as confident of that as I am of the weaker premise.)

Fifth: then such a world of stages would be exactly like our own simpliciter. There are no features of our world except those that supervene on the distribution of local qualities and their causal relations.

Sixth: then our own world is a world of stages. In particular, person-stages exist.

Seventh: but persons exist too, and persons (in most cases) are not person-stages. They last too long. Yet persons and person-stages, like tables and table-legs, do not occupy spatiotemporal regions twice over. That can only be because they are not distinct. They are part-identical; in other words, the person-stages are parts of the persons.

Let me try to forestall two misunderstandings. (1) When I say that persons are maximal R-interrelated aggregates of person-stages, I do *not* claim to be reducing "constructs" to "more basic entities". (Since I do not intend a reduction to the basic, I am free to say without circularity that person-stages are R-interrelated aggregates of shorter person-stages.) Similarly, I think it is an informative necessary truth that trains are maximal aggregates of cars interrelated by the ancestral of the relation of being coupled together (count the locomotive as a special kind of car). But I do not think of this as a reduction to the basic. Whatever "more basic" is supposed to mean, I don't think it means "smaller." (2) By a part, I just mean a subdivision. I do not mean a well-demarcated subdivision that figures as a unit in causal explanation. Those who give "part" a rich meaning along these lines<sup>4</sup> should take me to mean less by it than they do.

<sup>4</sup>Such as D. H. Mellor, in his *Real Time* (Cambridge: Cambridge University Press, 1981), chapter 8.