

Time Preference

18/2015

Political Dialogues

DOI: <http://dx.doi.org/10.12775/DP.2015.011>

Abstract

This paper has the following logical structure. First, I am trying to demonstrate how various theories of identity relate to the universal fact of time preference. Second, I will argue for the validity and superiority of psychological theory of identity to other theories and it will be shown that the psychological theory of identity can in principle explain the otherwise inexplicable brute universal fact of time preference. Third, after Nozick, the methodology is presented, in which it is claimed that when there are two distinct realms, the framework of one can serve as the *explicans* for the other. Thus, there is no danger of any logical fallacies, the most haunting one being vicious circle and even if the more fundamental realm is not the factually true explanation of the realm being in need of explanation it can greatly illuminate the latter. Finally, some implications of psychological theory of identity for the idea of time preference will be elucidated.

How does theory of identity correlate with time preference?

For the reasons that will become clear later, the only theory on which the universal fact of time preference could be

grounded is some reductionistic psychological theory of identity. What should be dismissed and what cannot account for time preference is any substance theory of identity. After all, time preference is all about discount rate and if any substance theory of identity is true, the fact of time preference would be totally inexplicable. But what exactly is time preference? Time preference basically means that for any given economic good, people universally prefer to get it sooner rather than later. The word of caution is due at this point. We need to conceive of an economic good rather carefully. *The same* good can denote different physical objects used to attain the same goal. And conversely: the same physical objects can constitute different economic goods for the simple reason that they can satisfy different needs.¹ For example, sunglasses at noon make a different economic good from sunglasses in the evening despite their identical physical make-up. The reason should be all too clear by now. Sunglasses at noon serve to prevent our eyes from being blinded by sunlight and in the evening we can wear them to look

¹ On the idea of economic goods see: Murray Rothbard *Man, Economy and State*, Ludwig von Mises Institute, 2011, p.410–416

neat at some fancy dress festival. We can value the fulfillment of those two needs differentially; so, sunglasses make two distinct economic goods depending on the time of the day. Consequently, and that is the crucial point, if we could obtain the sunglasses only on one occasion (for example we rent them and we are charged according to some hourly rate) we might prefer to have them in the evening- so later rather than in the morning. In conclusion, time preference operates only when it comes to the same economic goods. In other words, for a *given need* people universally prefer to satisfy it sooner rather than later.

If we want to somehow correlate a theory of identity with the fact of time preference, we can easily observe that any purely *substance* theory of identity wouldn't do. Such a theory can at best provide the purely formal condition for universal time preference, that is between A at t1 (giving a loan) and A at t2 getting it back with the interest rate, there must be the *relation of identity*; or, to put it more mildly there should be at least the belief at t1 that the recipient of the loan with the interest rate will be the same person. It is inconceivable for an individual A to lend some money and *believe* that it will be *somebody* else who will cash in the interest. So, it looks as if – to explain the universal fact of time preference – we are still limited by the idea (or a belief) of identity, that is, however distant the psychological relation is, an individual at an earlier time must believe that at t2 there still will be the same person (numerically). Yet, more importantly, there is another condition which, when complemented with the previous condition, constitute the sufficient condition for time preference to occur. That condition is Parfitian *Relation R*, which constitutes what matters alone in survival.

Relation-R is reducible to *psychological connectedness*, which Parfit (somehow arbitrarily) defines as some critical number of strong psychological connections, and to *psychological continuity*.² And obviously, psychological connectedness is what matters most to our problem of universal preference, and luckily this is at the same time what Parfit hails as “what is the most important both in theory and practice”³. Psychological connectedness varies to some degrees limited by the loss of identity and so does time preference. Low time preference cannot be extended *ad infinitum*. One cannot plausibly lend some money and agree to be repaid the loan after 200 years as there is simply no chance that there would still be *him* in 200 years. Let's illustrate some of the above points. Let's take a given individual or, speaking more precisely, a given body A at time t1. ⁴ If our theory of identity predicts that, say, from t1 to t2 body A is inhabited by the same person but at any times later than t2 it's definitely somebody else because, for example, the brain of that body was deprogrammed, then such a theory would have only marginal explanatory power. It would suc-

2 D. Parfit, *Reasons and Persons*, Oxford University Press, Oxford 1986, p. 206. Psychological connections hold between A at t1 and A at t2 when A at t2 has a memory of the experience he or she had at t1 and according to Parfit there is strong psychological connection between A and A1 at two distinct times when there is at least 50% of psychological connections present in ordinary day-by-day survival. On the other hand, psychological continuity additionally assumes that there is an overlapping chain of strong psychological connections.

3 D.Parfit, *Reasons...* p. 206

4 I assume that we do have at our disposal a criterion to predicate whether a given body is or is not the same over time. The criterion can be spatio-temporal continuity. That is if we can make sense of an individual body and we can trace it continually along spatio-temporal trajectory, we can be sure it is the same body. To put it differently, a body doesn't make leaps either in space or time.

ceed in explaining the infinite interest rate in the case of A1 lending money to B1 at t_1 and supposed to get the money back (plus interest rate) at t_3 , which is later than t_2 , when A ceases to exist. Yet, such a theory would fail to explain any time preference within the spectrum $\langle t_1, t_2 \rangle$.⁵ For example, if I lend 100\$ to the person B at t_1 and I am supposed to be returned the loan at t_2 when the recipient at t_2 is *equally*⁶ me (after all we are now considering the non-reductionistic theories of identity), such a theory cannot reflect the preference of the present over the future and thus fails to account for time preference. Yet, obviously we cannot simply assume the reductionistic psychological theory of identity is true and thus provides the necessary foundation for time preference. That would be a major methodological flaw. The point is to show that psychological theory of identity stands on its own and to show that *if* this theory of identity is true (or the most reliably accounting for our intuitions about our survival), the universal fact of time preference logically follows.

5 It seems to be the only reasonable solution. If I lend 100\$ at t_1 until t_3 when I know that I am going to be deprogrammed and so there is no me at t_3 (the psychological connectedness between the person at t_1 and t_3 is after all zero). Therefore, I would logically charge infinitely much as the recipient of the money I lent wouldn't be me but somebody else. Practically speaking, there would be no loan in the first place.

6 Of course, identity is an either-or relation; yet, if the relation is irreducible, no questions of psychological connectedness, which is after all gradable, can arise. After all, it is only within psychological account that one is able to explain the fact that even though A and A1 and A2 etc. are the same individuals over time, A would somehow identify less with the individuals more distant in time.

Arguing for the psychological theory

We can first endorse the psychological theory of identity negatively, that is by eliminating other theories. Let's first deal with the soul theory of identity. It can be easily noticed that this theory exemplifies substance theories. It basically claims that a given person is the same person over time when he or she preserves the same soul. There are some major philosophical problems that this solution runs into. First and foremost, it should be borne in mind that soul can be (and indeed usually is) conceived of as the featureless entity, regardless of whether it is considered "the vitality that distinguishes human beings (...) an ethereal but nonetheless physical entity" or an apparently "spiritual substance which is the subject of his mental states and the bearer of his personal identity".⁷ . To put it more precisely, soul cannot be detected by any empirical means, and more crucially, *that* soul or *this* cannot make any introspective difference. Now, let's imagine two persons: Greg and Erik. *Ex hypothesi*, neither is a zombie; they are both gracefully endowed with souls. Let's suppose Greg has the soul X and Erik has the soul Y. One day their souls swap places: soul X (until today inhabiting Greg's body) breathes life into Erik's body and soul Y breathes life into Greg.⁸ What has really happened on that day?

7 A. Quinton (1972) *The Soul* in J. Perry *Personal Identity*, University of California Press, Berkeley p. 53–72

8 On the original considerations about the irrelevance of the soul for persona identity see: A. Quinton (1972) *The Soul* in J. Perry *Personal Identity*, University of California Press, Berkeley p. 53–72. The reader is invited to bear with me and suspend the disbelief that souls can somehow travel—especially when they are supposed to be featureless blobs. Still, of course, the objection is valid whether an idea of movement can apply to immaterial objects.

Psychologically or empirically speaking, nothing. Neither Greg (as distinguished by the bodily criterion), nor Erik can spot any difference. On that day, they still remember their life before the soul travel. Not only there is no difference detected by them but also their respective friends cannot notice any change. Greg and Erik (still as distinguished by the bodily criterion) seem to display their usual behaviour and they have their respective memories intact. This thought experiment clearly shows that such an idea of a soul is simply empty. If soul-swap cannot make difference to its host, we should get rid of the idea of soul as a sole criterion of personal identity.

The second objection to soul is the following.⁹ Soul is a non-reducible *substance* and thus cannot be reduced to any psychological relations. Positively speaking, the soul can be considered a sort of *subject* or an immaterial entity being then necessary foundation for mental experience or spiritual élan vital breathing life into the physical body. Then, the argument against is that it is hard to imagine that the idea of numerical identity can apply to ideal objects. Of course souls must be somehow numerically distinguished if they are to account for the *separateness of persons* on their own. But if souls breathe life into limbs, we can imagine there is a possibility that souls are qualitatively identical (let's refer to the case of twinning, where the identical twin appears after the ensoul-

9 It is irrelevant whether we conceive of soul as élan vital or as Cartesian ego. The following critique applies with the same force to either of these concepts. The point is that any purely ideal objects which are qualitatively identical cannot be thought of as numerically distinct for the reasons that are explained in the text. On the various concepts of soul see: Anthony Quinton (1972) *The Soul* in John Perry *Personal Identity*, University of California Press, Berkeley p. 53–72.

ment of the zygote). We can assume then that the moving spiritual force is the same in case of those two twins. So then, to account now for the separateness of those two twins, we need to postulate that their souls are numerically distinct. But how can the same ideal object(s) be numerically different? Such ideal objects as souls or angels by definition cannot have any material manifestations. The analogy can come in handy. When, for example, there are two Windows 10 users, A and B. They can rightly claim them that they have two different Windows 10 insofar as they are installed into two different pieces of hardware. It's like having the same idea recorded on two different media. But in the case of soul there is no physical media by definition. Then, this idea of soul as the criterion of identity fails. Obviously, it does not imply that souls do not exist at all.

Incidentally, the same objection cannot be raised against Parfitian theory of identity. Quinton (who calls it an empirical soul theory) rightly observes that a psychological theory correctly accounts for the separateness of persons for even if two identical twins are born and they travel the world hand in hand, they still see the world from different angles. Moreover, he notes that they cannot occupy the same place at the same time¹⁰ So, the series of mental events or the chain of overlapping memories is *unique* even in the case of practically inseparable identical twins.

The second type of theory is the Embodied Mind Account by McMahan.¹¹ Here my method will be not to argue directly against it but to demonstrate the relative advantages of Parfitian theory

10 A. Quinton (1972) *The Soul* ... p. 58–61.

11 J. McMahan *The Ethics of Killing*, Oxford University Press, 2012, p.66–88.

over McMahanian one. First, Embodied Mind Account implies that when a given person collapses into Alzheimer, the relation of identity between P1 (a pre-Alzheimer person) and the present patient P2 (with advanced Alzheimer) still holds because there is still the same brain whose functionality is still preserved.¹² Yet, there are no strong psychological connections between P2 and P1. There is, let's suppose, not a single experience (or a belief, desire for that matter) that P2 remembers that P1 had at t1. So, for McMahan, P2 and P1 are the same person merely because there is the same functional brain. Furthermore, McMahan explicitly states that a person would survive being deprogrammed if there is some part of the original brain preserved still having the disposition to give rise to consciousness.¹³ In passing, we can notice that such a theory would fail to explain the universal fact of time preference. The horizon of low time preference cannot exceed the boundaries of our psychological connections. In other words, A at t1 would not go for any financial transaction if deprogrammed A at t2 (in McMahan's sense of identity) would not have a single trace of memory of the event at t1. Still, that cannot undermine McMahan's theory because it would be like putting the cart before the horse. What can be safely said is that Parfitian theory is simpler and more elegant being devoid of the insistence on the same brain. Furthermore, it accounts for our intuitions about survival, I believe, better than McMahan's. It is hardly believ-

able that the very same brain, even if all the psychological connections are lost, with its ability to give rise to consciousness is something that we can egoistically care about.

Now, it's high time to deal with Parfit's account of identity but it takes some conceptual preparation first. Parfit explicitly states that identity is a non-gradable relation. Parfit complements it with adding that survival is gradable (above some threshold, which shall be explained later) and what matters is survival. Parfit defines identity as a conjunction of one-to-one relation and psychological continuity and/or psychological connectedness (Relation R). Therefore, practically speaking, abstracting from some fancy thought experiments, Relation R in the real world guarantees identity simply because in the real world persons are not replicated yet and Division is still just a thought experiment.¹⁴ Then, if the relation is 1-to-1 (and it always is in the real world) then what matters is identity. So, however much Parfit concentrated on what matters and discarded identity as irrelevant, we can identify what matters (that is survival) with identity in our actual world.

Now, to the point. The relative advantage of Parfitian theory over *any* Physical Accounts¹⁵ is exemplified in the following imaginary scenario. I believe it counts much in favour of psychological accounts and it shows that it is not the same brain but rather psychological continuity that matters. Suppose, you've

12 J. McMahan *The Ethics...* p.66–67.

13 J. McMahan *The Ethics...* p.68. This position is still different from the belief that we are our bodies. Of course, the latter position can be easily debunked by saying that after all we are not our corpses, so the identification of persons with bodies fails.

14 D. Parfit *Reasons...* p. 200–302. Division is a thought experiment in which the author imagines the scenario when the two symmetrical hemispheres of the original persons are respectively grafted into his two identical siblings.

15 Among others, J. McMahan and T. Nagel with their insistence on some critical level of the sameness of the brain stand out as probably the most representative.

had the totally exhaustive time at work the other day and you come home being on your last legs about to fall asleep. When you already sleep tight, your roommate, who happens to be a neurosurgeon surreptitiously applies anesthetics to you and performs a surgery on you. He removes your biological brain and substitutes the silicon brain for it. The silicon brain has the same structural properties as your biological brain. If functionalism is true, you should wake up being psychologically continuous with the individual (P1) who was on his last legs after work the day before.¹⁶ So, in other words, you cannot notice any difference. The relation between P2 (the individual after the surgery) and P1 is absolutely the same as in normal (bodily continuity) survival. Would you conclude upon learning that now you have the silicon brain that the surgeon has committed a murder and the original person is dead? Would you mind as P1 the operation in the first place if you were promised that R-Relation is going to hold and only hardware (that is the brain) will be replaced? Of course not. I think it shows that what matters alone (especially when the relation is 1-to-1) is psychological continuity.

Nozick's explanatory framework

In his book *Anarchy State and Utopia*¹⁷, Nozick presents the possible ways of how two different realms can relate to each other explanatorily. Though Nozick considers explaining the political

in terms of non-political, we can apply his analyses to any apparently distinct realms and keep in mind, while quoting him, that our realm is the theory of identity and economics and especially time preference. So, Nozick presents the following schema: "The possible ways of understanding the political realm are as follows: 1) to fully explain it in terms of the nonpolitical; 2) to view it as emerging from the non-political but irreducible to it (...) 3) to view it as a completely autonomous realm. Since only the first promises full understanding of the whole political realm, it stands out as the most desirable theoretical alternative, to be abandoned only if known to be impossible". Then, Nozick distinguishes three types of defective potential explanations. "(...) a *law-defective* potential explanation is a potential explanation with a false lawlike statement and that a *fact-defective* explanation is a potential explanation with a false antecedent condition" Finally, Nozick identifies the *process-defective* explanation, which is an explanation which fails when the causal law it suggests is valid and yet it is the other antecedent A that is factually responsible for the consequence C.¹⁸ Let us try to apply Nozick's schema to our problem of explaining time preference in terms of the psychological theory of identity. What can be easily noticed is that our explanation of time preference can be *process-defective*, that is it may be genetic dispositions and environmentally acquired habits and beliefs that time preference could be fully reducible to. Still, if on the other hand, the said factors did not play a role when contributing to time preference, the psychological theory of identity would provide a very elegant explanation to the fact that the

16 D. Dennett, *Consciousness Explained*, Back Bay Books, 1992. Functionalism has it that the substance of the brain is irrelevant for the mental life, so we could in principle, implement the same mental life in different matter providing that its architecture/structural properties are identical.

17 R. Nozick, *Anarchy State and Utopia*, Blackwell Publishing, Oxford 2014, p.6-8.

18 Robert Nozick, *Anarchy...*p.7-8

more distant future we consider the less preferred *the same economic good* is. Let us take a closer look at the merits of our tentative explanation. First, the psychological theory of identity with its emphasis on the gradable nature of direct psychological connections kind of reflect the gradable nature of time preference. These two realms are at least structurally similar and they are at least partially positively correlated. That is the weaker the psychological connection is, the weaker the preference for the future (and the higher the preference for the present) that is the higher the interest rate charged by the loan-giver. Furthermore, the psychological theory of identity seems more fundamental than the fact of time preference; thus, the universal fact of time preference can be at least plausibly thought of as derived from the facts about our identity (or of what matters for our survival). Moreover, the psychological theory of identity is, trivially, psychological by nature and it may serve as the explanation of the psychological fact of time preference. So, it is conceivable that these two realms, though distinct, are not poles apart after all and what it may take to connect them is some bridge law. Finally, what is advantageous about explaining one realm in terms of another is the fact that in such cases, the logical fallacy of vicious circle is ruled out. The terms within the realm of identity are by definition distinct from the concepts from the realm of economics (and *a fortiori* time preference), so there cannot be the case that what is to explained is assumed in the *explicans*. In conclusion, although it may be the case that there were other factors factually giving rise to the universal fact of time preference, the facts about our identity (as postulated by the psychological theory of identity) might have equally well done so.

That observation means that psychological theory of identity, however plausibly false, sheds some light on the problem of time preference.

What does the psychological theory of identity imply for time preference?

In this section, I will try to examine what specifically follows from the assumption of the psychological theory of identity for the idea of time preference. First, a slight concession should be made. It is not case that our psychological theory of identity exhaustively does justice to the idea of time preference. That theory of identity can only account for the *universal fact of time preference* but it alone is unable to account for *individual time preference scales*. The factors contributing to individual time preferences are many and they may include: temperament, patience/impatience, the current financial status (it is easier to be a capitalist, that is the one who resigns from current consumption and instead invests to get the interest later, when one has already accumulated quite considerable wealth) etc. Yet, psychological theory of identity (or more precisely a belief that what matters for our survival are psychological connections, which are gradable) is enormously important and it does explain the universal fact of time preference. What is more, the belief that psychological connection matter is a contributing factor to individual time preferences. Let's imagine two scenarios with the same person P. It is the same person because, *ceteris paribus*, what changes is the person's belief about the strength of the psychological connection between P and P1 at t1 in the first scenario and between P and P1 at t1 in the second scenario. Let's say that the person persists in both scenarios because

the Relation R still holds (the psychological continuity condition is not violated) and that is the very same person across those two considered worlds. In the first scenario, P suffers a premonition that he can contract a neurologically debilitating disease, which would gradually deprive him of his memories as time goes by. In the second, no such premonition is felt and the person happily believes that he is going to have a strong psychological connection between P at t and P1 at t1. How would those two beliefs about psychological connectedness correspond to respective time preferences? Everything else equal, P in the former scenario would have much higher time preference than in the latter. It is merely because the degree of his psychological identification with *himself* in the future is much weaker in the first case than in the second one. Practically speaking, if P lends 100\$ in the first scenario, he is going to charge much higher interest rates than in the other and when *ceteris paribus* condition is met it is explained by the differential psychological identification with oneself in the future *alone*. In conclusion, psychological theory of identity accounts for the universal fact of time preference as well as it is a contributing factor co-determining the individual time preference scales.

Second of all, although, as mentioned, psychological theory of identity alone cannot provide the explanation for individual time preference scales, it can correctly explain the limits of time preference. Any loan given by A to B at t1 wouldn't be given for such a long time (terminating at t2), when if at t2 A wouldn't remember the very act of giving a loan to B. For A to give B a loan at t1, A should at least believe that he will remember the transaction at t1. It is the belief that there will be the psychological

connection between A and A1 that is sufficient for the very transaction to occur. Of course, it does not explain the specific interest rate charged by A for it is a multivariable function. If, say, it is soul that matters for survival, A would be prone to have infinitely low time preference- especially being rich already.

Conclusions

It appears that our agenda was of the logical *modus ponens* form. If a reductionistic psychological theory of identity is true, the universal time preference follows. If we can plausibly argue for the psychological account of identity, then time preference is true or plausible (depending on the validity of the psychological theory of identity). In other words, the universal fact of time preference is explicable only to that extent to which the psychological theory of identity is valid. Finally, if the psychological theory of identity is true, it can at least explain the formal properties of time preference, such as its limitations and its universality.

Bibliography

- Dennett D., *Consciousness Explained*, Back Bay Books, 1992.
- McMahan J., *The Ethics of Killing*, Oxford University Press, 2012.
- Nozick R., *Anarchy State and Utopia*, Blackwell Publishing, Oxford 2014.
- Parfit D., *Reasons and Persons*, Oxford University Press, Oxford 1986.
- Quinton A., *The Soul in J. Perry Personal Identity*, University of California Press, Berkeley 1972.
- Rothbard M., *Man, Economy and State*, Ludwig von Mises Institute, 2011.