

A Survey of Plausibility in Personal Identity Theory

G.J. O'Farrell

Abstract: In the attempt to find a plausible theory of personal identity, different kinds of concerns present different sets of problems. The search is further complicated by the implications of different ontological foundations in philosophy of mind. Other-regarding concerns such as numerical identity over time and moral accountability, as well as self-regarding concerns such as self-identification and personal character can be approached by a variety of epistemological theories. Complimentary theories which are insufficient on their own can be combined to more rigorously answer certain other-regarding or self-regarding concerns, but fail to address both sets of concerns simultaneously. While these sets of theories have practical value in our daily lives, they fall short of providing a metaphysical theory of personal identity. The popular personal identity theories discussed are dominated by materialist or dualist ontologies, and may provide answers to epistemic questions, but a unified metaphysical theory of personal identity appears to require a different approach which challenges popular ontological assumptions.

Many philosophers have understandably struggled with the question of what constitutes personal identity. It is a complex topic which intersects with most, if not all aspects of human

life. Having become acquainted with several popular but divergent theories of personal identity, this paper will discuss which, if any, may be the most plausible. In this discussion, certain ontological assumptions about the mind, the body, and the world will carry different implications. The theories that I am familiar with are generally based in either a materialist or dualist ontology. For example, the body criterion for personal identity, which states that identity over time is established by a continuous physical body, is rooted in a materialist ontology, while the soul criterion, which states that identity over time is established by a continuous soul, relies on dualism.¹ Because my preliminary study of personal identity theories aims to understand the phenomenon by way of its relevance to particular questions and concerns, I will argue (1) that different sets of epistemological theories can be combined to successfully address either self-regarding or other-regarding concerns, which is to say that they help us to obtain genuine knowledge in response to particular sets of questions, (2) that idealism is a useful and worthwhile ontological point of view to consider when trying to understand self-regarding concerns of personal identity, and (3) that a metaphysical theory of personal identity requires a different approach.

First let us consider other-regarding questions of personal identity, specifically numerical identification and accountability across time. When it comes to identifying an individual and holding them responsible for past behaviour, I argue that the most plausible theory of personal identity includes the aforementioned biological criterion, which can be described in the following way: If X is a person at t1, and Y exists at any other time, then X=Y if and only if Y's biological organism is continuous with X's biological organism. However, this would be insufficient in scenarios where there is no way to confirm the continuity between X and Y. We

¹ David Shoemaker, *Personal Identity and Ethics: A Brief Introduction* (Peterborough: Broadview Press, 2009), pp.24-29.

could plausibly make a mistake and identify X at t1 with their twin, Y at t2, believing them to be continuous because of their biological similarity, even though Y is in this case obviously not responsible for X's behaviour, both of them being numerically distinct. To account for this variable we can add to the argument the psychological criterion for personal identity, which can be expressed as follows: X at t1 is the same person as Y at t2 if and only if Y is uniquely psychologically continuous with X. The psychological criterion on its own is insufficient in establishing numerical identity across time because a person's psychology is immaterial, and there is no way an outside observer can get inside another person's mind to confirm psychological continuity. However, I would claim that it is a necessary addition to the body criterion if we want to avoid cases of mistaken identity, as in the twin case just mentioned, as well as in determining the degree of accountability, as I will illustrate shortly. Combining these two criteria result in the following argument: If X is a person at t1, and Y exists at any other time, then X=Y if and only if Y's biological organism is physically and psychologically continuous with X's biological organism, and Y is therefor morally accountable for X in proportion to the extent that Y and X are psychologically continuous.¹

To illustrate the second half of this argument, consider subject A, who commits a crime and then suffers an injury resulting in retrograde amnesia, erasing their memory up until after the crime was committed. It seems absurd to say that, because the person who committed the crime is not psychologically continuous with A, that they are no longer the same person, even though A may have a hard time associating with their own past actions. Relying only on the psychological criterion in this scenario would imply that after suffering amnesia, the person who committed the crime ceased to exist, and A spontaneously appeared. However, according to the psychological criterion, if amnesiacs retain just a small portion of their memory and learned behaviour, this

¹ Shoemaker, *Personal Identity and Ethics*, pp. 60-77.

limited psychological continuity would imply that A is still the same person, but to what extent? Based on A's psychology alone, not even A would be able to self-identify across time as the criminal. This is just one example which shows that psychological continuity alone is insufficient for numerical re-identification, and at the same time illustrates how the biological criterion of identity accommodates a psychological relation. We know that A's psychological discontinuity is caused by physical changes in A's brain, which is biologically continuous with A prior to suffering from amnesia. Biological continuity is both necessary and sufficient for identifying A as the criminal, but in practice, when determining the extent to which A is morally accountable for the crime, we would tend to give special consideration to the degree of psychological continuity between the amnesiac and the criminal. Because of A's psychological discontinuity, rather than being imprisoned, they may be put in a psychiatric hospital to help them regain lost memories. If we imagine that another person, subject B, were to confess to the crime, claiming that they vividly remember it, this would not be sufficient for identifying B as the criminal. We would require new evidence indicating that not only is B psychologically continuous, but biologically continuous with the criminal as well, and that somehow we have mistakenly identified A. Otherwise we would just say that B is deluded. However, if it were the case that we were mistaken, it follows logically that B would be morally accountable to a greater extent than A would be, because B has stronger psychological continuity with the criminal. Rather than being admitted to a psychiatric institution, B would face the full extent of the law for their crime.

I have given one example to illustrate how, when it comes to certain practical other-regarding concerns of numerical identification and moral accountability across time, both the biological criterion and the psychological criterion work together to form a plausible and useful theory. However, when it comes to self-regarding questions and concerns, such as self-identification and characterization, this theory runs into trouble in two ways. First, it is unable to

fully account for the explanatory gap between material events and subjective experience. The explanatory gap problem essentially asserts that a materialist ontology fails to account for the qualitative experience of consciousness.¹ While it is easy to concede that I am a biological organism with some degree of psychological continuity, this doesn't explain the subjective feeling of being myself and nobody else. What does the biological criterion say about consciousness, or character? The revised biological criterion, above, may be perfectly useful in numerical identification, but to me, the matter of my identity is not purely quantitative, it is a qualitative phenomenon. The second problem, related to the first, is that it is unclear what conditions a biological organism requires to be a person, or to what degree it can be altered or destroyed while maintaining the same identity. If I am essentially a biological organism, I can logically claim that I was a fetus, but there appears to be no clear way I can determine the point at which that fetus became a person, and even with the addition of the psychological criterion, I have no sense of continuity between my mind now and that infant. If I lose control of my body by slipping into a coma, the biological criterion would seem to imply that my personal identity persists, despite my having no conscious experience. Or, presuming that some conscious experience continues to occur in my brain, the biological criterion would then imply that identity can be contained entirely within the brain even when I have no control over the rest of my body. Similarly, there are numerous reports of out-of-body experiences in which people feel they have left their body, although the brain is still active.² Ultimately, in a materialist view of the world, these scenarios appear to present questions with no clear rational solution, but it is possible that another kind of personal identity theory can accommodate these concerns.

When it comes to these self-regarding questions and concerns, I argue that the most

¹ David Chalmers, "Facing up to the Problem of Consciousness" in *Journal of Consciousness Studies*, Vol.2 No.3 (California: Imprint Academic, 1995) pp.200-202.

² Shoemaker, *Personal Identity and Ethics*, pp.79-83.

plausible theory of personal identity is yet another combination of complimentary theories. First, the narrative criterion of identity: What makes an action, experience, or psychological characteristic properly attributable to some person (and thus a proper part of their identity) is its correct incorporation into the self-told story of their life,¹ and second, philosopher Derek Parfit's theory that identity doesn't matter (IDM), but psychological continuity does.² I find these two theories particularly complimentary within an idealistic ontology. Idealism avoids a number of problems by claiming that there is no attainable genuine knowledge of a mind-independent material world, and that what we perceive as such is easily reducible to mental events. In short, we do not have access to a material world, and therefore cannot authoritatively claim that a material world exists. All we really have access to are our own subjective experiences. My own mental states are, to me, as real as anything can possibly be. I can never attain genuine knowledge of a reality beyond my own subjective experience. This perspective is generally discarded as too extreme or unusual, but as philosopher of mind John Heil points out, it is difficult to imagine how it could be proven to be false.³ It may be unpopular to appeal to this ontological view, but I would argue it is parsimonious in the present discussion.

The narrative criterion of identity works well for an idealist, and also helps to solve the problems of materialist theories. Narrative identity does not depend on any material body or substance, but rather on the subjective narrative that I create to connect all of my various experiences into one continuous story. The explanatory gap problem is no longer an issue here, since there is no attempt to reduce mental events to material causes. Instead, it is my subjective experience, rather than a biological organism, which provides the necessary conditions for

¹ Shoemaker, *Personal Identity and Ethics*, pp. 87-90.

² Derek Parfit, *Reasons and Persons*, (Oxford: Clarendon Press, 1984), pp.213-217, accessed November 24, 2014, http://commonweb.unifr.ch/artsdean/pub/gestens/f/as/files/4610/17613_101712.pdf.

³ John Heil, *Philosophy of Mind: A Contemporary Introduction*, 3rd ed. (New York: Routledge, 2013) pp. 38-40.

personal identity. The material world, including my body, is just another subjective immaterial experience, which I can incorporate into the narrative of my identity. This doesn't make the narrative criterion completely incompatible with materialism, because all one has to do is correctly incorporate materialism into one's narrative, in that one would consider subjective experiences to be corresponding with an external material world, and reducible to material causes, although the explanatory gap would still be a problem.

The problem of degree mentioned earlier, of determining at what point and to what extent my personal identity can be considered to exist, is also solved by the narrative criterion. I no longer have to wonder to what degree my physical body must develop or remain intact in order to preserve my identity. Instead, my identity is maintained so long as I incorporate my experiences into my own narrative correctly.⁴ The question then naturally arises, what would be the correct way to construct my narrative identity, and is there an incorrect way? I believe these questions can be answered by incorporating Parfit's IDM theory to complement a narrative theory of identity. Parfit claims that what really matters is not identity, but psychological continuity, illustrated with compelling thought experiments, which I won't reiterate here.² By elevating the importance of psychological continuity over personal identity, we can answer the question about how to correctly incorporate experiences into one's narrative. When applied to narrative identity theory, Parfit's claim implies that the correct way is a psychologically continuous way, and that a psychologically discontinuous narrative would be incorrect, which is to say, it is possible to mistakenly identify myself. This may be counter-intuitive at first, but it has practical value in our lives. If someone spontaneously starts to identify themselves as the reincarnation of Jesus Christ, or Socrates, we would say they are no longer themselves. We

¹ Shoemaker, *Personal Identity and Ethics*, pp. 91-93.

² Derek Parfit, "Personal Identity," in *Personal Identity*, ed. John Perry (California: UC Press, 2008) pp. 199-220, originally published in *The Philosophical Review*, vol. 80, no. 1 (January 1971).

would also characterize the retrograde amnesiac described in the earlier example in the same way. The break in psychological continuity would constitute a broken sense of identity. The construction of a narrative in a psychologically discontinuous way creates a pathological identity which we would only consider rehabilitated if psychological continuity could be restored to the narrative.

Taken together, these theories can be expressed in the following way: What makes an action, experience, or psychological characteristic a proper part of the personal identity of X is its psychologically continuous incorporation into the self-told story of X's life. This helps us deal with self-regarding concerns where neither the biological or psychological criteria can, and where the individual theories of narrative identity and IDM are insufficient on their own. I could now answer the question of "Who am I?", or explain my individual characteristics, as we naturally tend to do in conversation, by referring to my personal narrative. If I lose the use of my limbs or slip into a coma, my identity won't suffer so long as I incorporate these experiences in a psychologically continuous way. However, there are still problems with this theory.¹ First, the end-points problem claims that narrative identity fails to provide a clear beginning and end to the existence of persons. Advocates of narrative identity may argue that the theory assumes a biological criterion of numerical identity, implying that the end points are located at the beginning and end of the existence of the biological organism, but this simply isn't the case. Narrative identity allows us to logically include events outside of our biological existence, such as being born, or even conceived, or to anticipate events which we cannot subjectively experience, such as the treatment of our bodies after death. For example, I learned as a child that my conception was not intended or planned, and this has become a part of my narrative identity that influences my attitudes and behaviour. Narrative identity also allows me to rationally

¹ Shoemaker, *Personal Identity and Ethics*, pp. 96-102.

anticipate how my body will be treated and how my life will be mourned or remembered when I am dead, implying that identity can extend beyond the numerical confines of my existence as a biological organism. This illustrates how, even though the biological criterion is useful, it doesn't address the same set of concerns that the narrative criterion does, and does not solve the end-points problem. All of this suggests that the narrative theory is not really concerned with personal identity at all, but if we recall and endorse Parfit's claim that identity doesn't matter, then we can say that this isn't really a problem. I may not be numerically identical with an embryo or a corpse, but I can still incorporate these parts of my existence into my narrative in a psychologically continuous way, and when it comes to self-regarding concerns, that psychological continuity is what really matters. However, the implication here, that identity is not exactly personal, and that it has no clear beginning or end, is unsettling.¹

The second problem which is raised in response to narrative identity is that it is unclear whether the theory describes how identity really is, or prescribes how it should be, however I would argue that this criticism misinterprets narrative identity theory. This counter-argument relies on the assertion that some people simply do not identify themselves as existing within a cohesive narrative, but rather experience their lives as isolated episodes, distinct and disconnected from each other. My claim is that this is still a valid form of narrative identity, and that it is only confusing insofar as it is unusual. Here Parfit's emphasis on the importance of psychological continuity is useful. Consider an analogy: *The Watchmen* is a popular graphic novel that was adapted into a feature film. The narrative is essentially the same in both mediums, or in other words, both mediums maintain essentially the same narrative continuity. However, as a graphic novel it takes the form of distinct episodes, separated by panels and pages. Each panel can be viewed in isolation as a distinct event, or with greater or lesser degrees of relation to other

¹ Shoemaker, *Personal Identity and Ethics*, pp. 110-112.

panels in the novel. A visual artist may very well examine a graphic novel in this way. As a feature film, many of the details of the graphic novel are deemed to be less important to the narrative, and so are left out during production. The film is recorded in frames, another kind of panel, which are played back so fast that we never perceive any separation from one frame to the next, resulting in a single continuous narrative. Like two individuals who construct their narrative identities in different ways, both methods of narration are equally valid for communicating the story of *The Watchmen*, carrying with them different advantages or disadvantages. The graphic novel is able to include more detail, while the feature film is able to streamline information to focus on what is most important. Whether I perceive the narrative in one form or another is irrelevant. Both are constructed with distinct events and experiences, and the different ways individuals process these events does not make narrative identity implausible. In other words, I reject the claim that a particular form of narrative continuity is necessary to make narrative identity theory intelligible. I argue that instead, all that narrative identity requires, at least when combined with Parfit's IDM theory, is self-identification in a psychologically continuous context.

The last criticism of narrative identity is already implied in the argument of this paper: It is not at all useful as an epistemological theory when it comes to other-regarding questions and concerns, such as numerical re-identification or accountability. There are other potential problems with narrative identity, and although I believe these may be solved by the incorporation of an idealistic ontology, that is too broad a topic to be tackled here. So what we are left with in this discussion are two sets of concerns, self-regarding and other-regarding, which can each be addressed with two sets of theories for personal identity, one of which includes the claim that identity doesn't really matter, and this brings me to my final point. I have tried to answer the question: What theory of personal identity is the most plausible? I have argued that no single

theory can plausibly address all of our questions and concerns, but that different combinations of theories can successfully address different sets of questions. Assuming for the moment that my argument so far is sound, there is one major question that remains unanswered: What is personal identity? I can not agree confidently with Parfit's claim that identity does not really matter unless I can claim to know what identity is. So far I have dealt with epistemological theories which describe how we can attain genuine knowledge about a person's identity in response to particular questions, but I have completely ignored discussion of a metaphysical theory which describes what personal identity actually is. For now, this kind of pragmatic discussion of a topic as complex as personal identity is as much as I can aspire to. It is possible that personal identity is just too broad and dynamic a phenomenon to be explained by a single unified theory, and consequently requires multiple explanations for multiple applications. I am not qualified or knowledgeable enough to propose a metaphysical theory of identity. What I will tentatively suggest, however, is that in order to arrive at an adequate metaphysical theory of identity, we may have to begin with an idealistic ontology, pay closer attention to the role of empathy and relationships in the way we construct our own narratives, and re-frame questions of personal identity appropriately. If narrative identity and IDM theories imply that identity can include parts of life we cannot recall but which are relayed to us, events which occur after death and which involve other people, this would seem to blur the boundary between self-regarding and other-regarding concerns, in that questions like "Who am I?" can only be answered by referring to both personal and inter-personal experiences. Perhaps when seeking a metaphysical theory, instead of asking "What is personal identity?" we may want to first ask "Is identity personal?"

Bibliography

1. Chalmers, David. "Facing up to the Problem of Consciousness," *Journal of Consciousness Studies* Vol.2 No.3: pp. 200-19. California: Imprint Academic, 1995.
2. Heil, John. *Philosophy of Mind: A Contemporary Introduction*, 3rd ed. New York: Routledge, 2013.
3. Parfit, Derek. *Reasons and Persons*. Oxford: Clarendon Press, 1984. Accessed November 24, 2014. http://commonweb.unifr.ch/artsdean/pub/gestens/f/as/files/4610/17613_101712.pdf.
4. Parfit, Derek. "Personal Identity," in *Personal Identity*, ed. John Perry. California: UC Press, 2008. Originally published in *The Philosophical Review* Vol.80 No.1, January 1971.
5. Shoemaker, David. *Personal Identity and Ethics: A Brief Introduction*. Peterborough: Broadview Press, 2009.