

*The Physical, the Subjective, and the Social  
in the Debate on Personal Identity:  
A Critical and Historical Overview  
of Neo-Lockean Accounts*

*Lo físico, lo subjetivo y lo social en el debate  
sobre la identidad personal: Una visión crítica e  
histórica de las teorías neolockeanas*

ALFONSO MUÑOZ-CORCUERA

Complutense University of Madrid, Spain  
alfonso.m.corcuera@gmail.com

DOI: <https://doi.org/10.15366/bp2021.28.001>  
Bajo Palabra. II Época. N°28. Pgs: 21-46  
Orcid: <https://orcid.org/0000-0003-0098-1489>



*Recibido: 30/09/2021*  
*Aceptado: 12/11/2021*

This research was funded by the Ministry of Science and Innovation of the Spanish Government through the Research Project "Institution and Constitution of Individuality: Ontological, Social and Legal Aspects" (PID2020-117413GA-I00 / AEI / 10.13039/501100011033).



## Abstract

There are, at least, three fundamental dimensions in what we all are: a physical, a mental and a social dimension. In this paper I will focus in how, in the history of the debate about personal identity, different neo-Lockean theories have tried to account for these dimensions. The main goal will be to show that the only neo-Lockean theory that can account for these dimensions is what I call the New Narrative Approach, that establishes a distinction between three entities that we are: human beings, selves, and persons.

*Keywords: Personal identity, neo-Lockean theories, History, Narrativism.*

## Resumen

Existen al menos tres dimensiones fundamentales en aquello que todos nosotros somos: una dimensión física, una mental y otra social. En este artículo me centraré en cómo, en la historia del debate sobre la identidad personal desde una postura neo-lockeana, se ha tratado de dar cuenta de las tres dimensiones. El objetivo será demostrar que la única teoría neolockeana que ha sido capaz de hacerlo es lo que denomino la Nueva Perspectiva Narrativa, caracterizada por establecer una distinción entre tres entidades que somos: seres humanos, yoés y personas.

*Keywords: Identidad personal, teorías neolockeanas, historia, narrativismo.*

## 1. Introduction

We are material beings. Or so it seems. The debate on personal identity mostly abandoned long ago the question of whether our identity depends on the existence of an immaterial soul or not. As early as the 17<sup>th</sup> century, Locke held that even if we had a soul, it would not determine our identity by itself, since what constitutes ourselves as a person (or a self)<sup>1</sup>—that is, as “a thinking intelligent being, that has reason and reflection, and can consider itself as itself, the same thinking thing, in different times and places” (Locke, 1975:335)—is our consciousness. If the same soul participated in the lives of two human beings, for example Socrates and your neighbour, but it did not carry with it Socrates’ consciousness to the body of the latter, then we would be dealing with two different persons. If on the other side, our soul were replaced by someone else’s soul while we are still alive, but our consciousness remained unaffected, then we would still be the same person (Locke, 1975:337-40).

Locke’s position has been very influential in the development of the debate on personal identity within analytical philosophy. Not only because of the refusal of the importance of the soul. But also because of the way the debate was framed. According to Flew, Locke’s contribution to the problem of personal identity was fourfold: first, he stressed the importance of the problem, linking it to questions of moral responsibility; second, he noted that thought experiments were relevant to the discussion; third, he specifically highlighted the fact that the identity of persons posits special problems, since contrary to what happens with material or immaterial objects, it does not depend on the unity of substance; and fourth and most important, he proposed his own account of the problem, which was the starting point of the debate within analytical philosophy (Flew, 1951:53). Not in vain all the accounts of the problem of personal identity that I will be dealing with in the following pages can be labelled as “neo-Lockean”, since all of them try to stand for one of the most basic of Locke’s intuitions:

- (1) We have to distinguish between at least two kinds of entities that we are, mental selves (or persons) and biological human beings, the former being understood not as pure egos but as series of interrelated mental events which are part of the same consciousness.

---

<sup>1</sup> Note that both terms were used as synonyms by Locke and most of neo-Lockeans.

Besides that, neo-Lockean theorists would also like to agree with Locke in two more points:

(2) that the identity of selves (or persons) is what matters to us when we care about survival and moral responsibility; and

(3) that the identity of selves (or persons) is determinate. That is, that there must always be a yes or no answer to questions regarding our identity as selves (or persons).

However, sometimes neo-Lockeans have found impossible to hold these two latter claims and they have given up one of them, or even both. But more on that later.

What I will present in this article is a history of the debate on personal identity within analytical philosophy since the mid-twentieth century to the present. Although obviously, as any historical account, it will be partial and slightly biased. First, because as I have just said, I will focus specifically on the neo-Lockean accounts of the problem, leaving aside other approaches as for example animalism (see Olson, 1997). Second, because it will be a simplified explanation which will ignore some of its details in favour of clarity and brevity. And third, because I write this history from a certain present point of view: I endorse a new branch of narrative theories of personal identity which holds that establishing a distinction between three entities that we are—human beings, selves, and persons—is the best way to give a full account of the problem of personal identity from a neo-Lockean perspective. In this sense, my presentation of the historical development of the problem is intended to prove that this new narrative approach is the one that makes more justice to Locke's original statements regarding the three claims stated above. To do this I will pay special attention to how the physical, subjective, and social dimensions of our existence fit together in the different accounts of what we are that have been offered from a neo-Lockean perspective once we have ruled out the importance of the soul.

Getting down to business, I will divide the history of the debate on personal identity within analytical philosophy into three stages. Taking some works as milestones, it can be said that the first stage, which I call "Objective Approach" (OA), started with Grice's discussion on Locke's theory in "Personal Identity" (Grice, 1941) and finished with the publication of Parfit's *Reasons and Persons* (1984). Regarding the second stage, which I call "Classic Narrative Approach" (CNA), I will assume that it started with MacIntyre's *After Virtue* (MacIntyre, 2007) and finished with Strawson's harsh criticisms in "Against Narrativity" (Strawson, 2008). And finally, with regard to the third stage, which I call "New Narrative Approach"

(NNA), we can consider that it started with Schechtman's reply to Strawson in "Stories, Lives and Basic Survival: A Refinement and Defense of the Narrative View" (Schechtman, 2007) and is still being developed.

Obviously, the distinction between these three stages is very rough, and both the distinction and the milestones could be the subject of many criticisms. Specially, the setting of definite endpoints to the first two stages is much controversial, since there are philosophers who are still working within the OA and the CNA. However, I consider that doing so after the publication of Parfit's and Strawson's works respectively is either redundant or it implies not having understood well the importance of these works, as I will try to show in the next sections.

## 2. The Objective Approach

Given that the accounts of the problem of personal identity I will be dealing with depart from Locke's theory, it will be worth saying a few things about it and the way it was interpreted by the first proponents of the OA. As I noted in the introduction, Locke held that personal identity depends on the continuity of *consciousness*. However, he was not completely clear about what he meant with this word, and so his view is subject to very different interpretations. For example, Flew notes that sometimes Locke seems to use "consciousness" as a synonym of "self-consciousness", while in other places he seems to be referring to "self-awareness". But in the end, he comes to the conclusion—and we can take Flew's interpretation as representative of the OA in this regard (see e.g. Grice, 1941:341; Williams, 1957:233)—that in his main statements Locke used "consciousness" as equivalent to "memory" (Flew, 1951:55). Locke indeed made emphasis on the importance of memory regarding identity statements, and in this sense, we can understand Flew's interpretation. For example, according to Locke, if a man loses the memory of some actions he did in the past, he would not be the same person who did those actions, even if he is the same human being (Locke, 1975:343-4). And although it could be interpreted that in this example Locke was taking memory only as an evidence of the continuity of consciousness and not as consciousness itself, or as being constitutive of its continuity, this interpretation was widely accepted within the OA.

We should note that none of the OA proponents were original in interpreting Locke in this way. Already in the 18<sup>th</sup> century Locke's theory was criticized in ways that imply the same kind of interpretation. For example, Butler stressed what is known as the *circularity problem*, that is, the fact that memory cannot constitute personal identity since remembering our past actions already presupposes our past

existence and thus personal identity (Butler, 2008:100). For his part Reid highlighted the *transitivity problem*, that is, the fact that transitivity is a property of the identity relation (if person A is the same as person B, and person B is the same as person C, then person A is necessarily the same as person C) but not of memory (if person A remembers himself doing what person B did, and person B remembers himself doing what person C did, it is not necessarily true that person A remembers himself doing what person C did) (Reid, 2008:114-5).

In any case, what is important for present purposes is that this interpretation of Locke's theory had two consequences for OA proponents. The first one is that they had a reputed philosopher holding an obviously wrong theory that they could amend and discuss. The second one is that in amending their memory interpretation of Locke's theory, they could ignore the subjective dimension of consciousness, since they understood memories as objective and isolated phenomena, just as the neural connections which encoded them. In this sense they understood memory (or psychological) continuity as an objective criterion of personal identity which they could oppose to another objective criterion, bodily continuity. Thus, we can see why I label this approach as the Objective Approach.

This objectivization of memories and other mental processes is the result of OA's endorsement of a wider ontological framework: *reductionism*. Parfit defines reductionism from the idea that someone is a reductionist if he believes that "a person's existence just consists in the existence of a brain and body, and the occurrence of a series of interrelated physical and mental events" (Parfit, 1984:211). But here it is important to note that, according to reductionism, these physical and mental events can be described in an *impersonal* way, that is, without claiming that they belong to, or are experienced by, a person (Parfit, 1984:210). Against this, Parfit thinks that non-reductionists have two options left: either considering that (a) we are separately existing entities, distinct from our brain and body, like for example Cartesian pure egos; or (b) that although we are not separately existing entities, personal identity consists in a further fact besides physical and/or psychological continuity (Parfit, 1984:210). However, later Parfit notes that one cannot coherently hold that personal identity consists in a further fact unless one also believes that we are separately existing entities (Parfit, 1984:240). In this sense, he presents reductionism as if it were just an innocent denial of Cartesian dualism, as it has been pointed out by Ricoeur (1994:131-2), McDowell (1997:244) and Shoemaker (1997:135), among others. But although Parfit does not explicitly say it, this kind of reductionism has further consequences. For example, the claim that we can describe persons' experiences in an impersonal way seems to imply a reduction of mentality to physical or functional terms as well, thus ignoring the subjective

dimension of experience, as I have already noted (Ricoeur, 1994:133; Shoemaker, 1997:139). It also seems to deny the relevance of the social aspects of our life, thus making it difficult to maintain that personal identity is still relevant to moral responsibility. In any case I will come back later to this.

Since OA proponents generally accepted reductionism—although perhaps not all of them in Parfit’s extreme form—they all thought about it as completely consistent with Locke’s account of personal identity. So, we can understand OA as the attempt to hold the three claims stated in the introduction along with a fourth claim:

- (4) Reductionism is true, so that the interrelated physical and mental events which constitute our personal identity can be described in an impersonal way, that is, without presupposing that they belong to, or are experienced by, a person.

Evidently, OA proponents differed in their opinions about the interrelations between the physical and mental events taking place in our body and brain which were necessary for personal identity. Nevertheless, most of them amended Locke’s theory in similar ways, so that their proposals were not subject to the problems raised by Butler and Reid. For example, most of OA proponents accepted Shoemaker’s concept of quasi-memory to avoid the circularity problem (Shoemaker, 1970:271).<sup>2</sup> And most of them also introduced in their accounts more psychological events besides memory (like beliefs and character traits) to avoid the transitivity problem as well (Parfit, 1984:204-5). In this sense, the main position within OA was that personal identity consisted in psychological continuity, this understood as an objective phenomenon. In any case, I will not say anything else about the different existing accounts of the problem of personal identity within the OA. Instead, I will present what we can label as the “*inconsistency problem*”, which affects all these accounts and thus can be considered as the main reason why OA failed to give an acceptable account of the problem: if reductionism is true, then personal identity is either indeterminate or unimportant, so that claim (4) is inconsistent with claims (2) and (3) at the same time.

---

<sup>2</sup> As Butler (2008) noted, if we have a memory of a past event it implies that we already existed when that event took place and that we were there to experience it. Shoemaker proposes the concept of quasi-memory to refer to past events independently of the person who experienced that event in the first place (Shoemaker, 1970:271). If the person who experienced that event is the same as the person remembering it, then we can talk of memories. But it could be possible that someone remembers an event which was experienced in the first place by someone else—think about *Blade Runner* or *Memento* (see Muñoz Corcuera, 2009). This person would not be remembering, but quasi-remembering a past event which was not experienced by himself. Thus, if we talk of quasi-memories instead of memories, we will be avoiding the circularity problem.

As it is well-known, thought experiments were the main methodological tool for OA proponents. The hypothetical cases which put their basic intuitions about personal identity to the test—brain transplants, mind transfers, complicated surgeries, teleportations, and so on—were highly influential. And as Parfit forcefully defended throughout his *Reasons and Persons*, if we hold reductionism then there will always be logically possible situations in which there would be no significant facts which would let us give an answer to questions about personal identity. No matter which version of the OA we hold. Perhaps the most famous cases among this kind of thought experiments are those involving the duplication of persons. That is, cases in which we have two future persons who seem to be right in claiming the identity of a single past person according to our identity criterion. That would be the case, for example, we could split somehow into two beings like an amoeba, so that each of these had the same physical and mental properties that we have at the moment, but only half of our original brain and body. In cases like this, in principle it would be impossible to say which of the future beings, if any, is identical with ourselves. Both of them would be physically and psychologically continuous with us, and so both could claim our identity. However, they would not be identical with one another, and since identity is a transitive relation, both could not be right.<sup>3</sup> Thus, given that there are no significant facts that could let us give an answer to questions about personal identity in this kind of cases, then in principle we must accept that personal identity is indeterminate. So, if we accept reductionism, we seem to be forced to give up claim (3) and accept that in some cases it will be impossible to say if we would survive or not.

On the other hand, there is an alternative to giving up claim (3). As Parfit notes, whenever our criterion of personal identity is unable to provide an answer to a certain case, we can extend our criterion so that we can give one (Parfit, 1984, p. 241). However, given that in the kind of cases we are dealing with there are no significant facts which we could consider to take a decision, any answer we could give would have to be based on trivial or external facts. For example, we could solve duplication cases by stating that whenever there is a duplication, the resulting being who is closer to a certain point will keep the identity of the original person, while the other one will not (*cfr.* Williams, 1957: 239). Or, as some OA proponents did, we can hold that personal identity consists in psychological continuity, except in the cases in which there are more than one person

---

<sup>3</sup> In this regard, Schechtman (1996:30) considers that duplication cases are just a particular case of the transitivity problem.



who is psychologically continuous with a past person, in which case none of them is the same as the past person (see Nozick, 1981:32-40). However, although this would let us keep claim (3), it would force us to abandon claim (2) in exchange, since whatever trivial or external fact we take as decisive in this regard, it will be unable to bear the importance we attribute to personal identity when we care about survival or moral responsibility<sup>4</sup> (Williams, 1970:178). Think about what I have labelled elsewhere as the “we-might-all-be-dead problem” (Muñoz-Corcuera, 2021b:82): if personal identity consists in psychological continuity, and duplication eliminates personal identity, maybe none of us is the same person who was yesterday because there is a mad scientist in Mars who has built exact replicas of ourselves who are psychologically continuous with our past existence; but if our personal identity has been eliminated just because of that, then personal identity cannot be what matters to us when we care about our survival, since we would not think that we have died just because there is a replica of ourselves somewhere in Mars.

Evidently the inconsistency problem does not force us to abandon the OA. Claims (2) and (3) are not undeniably true. Neo-Lockeans would like to hold them because they are intuitions that most of us have regarding our existence. But they are just that, intuitions. We tend to think that personal identity is what matters in survival because we think that what matters to us when we care for our future is whether *we* will be there in the future, living our lives and experiencing our experiences. In the same way, we tend to think that personal identity has to be determinate because it is difficult to imagine that, given a future situation in which a person will be tortured, sometimes it will be indeterminate whether we should be afraid of that situation because it is indeterminate whether we will be that person or not. But we could be wrong in both cases. And if one holds reductionism—and remember that only Cartesian dualists would dare not to do so!—that is the obvious choice. However, I find this conclusion highly unsatisfactory. First, because it seems to me that our intuitions in this regard are to be taken seriously. Who else, if not ourselves, can tell us what matters to us when we care about our survival? And second, because I do not see why we should accept reductionism. Parfit’s presentation is highly tendentious, as we have already seen. But if reductionism implies the denial of the subjective dimension of consciousness, we should consider that eliminativism is far from being the most popular position within contemporary philosophy of mind.

---

<sup>4</sup> This was Parfit’s choice, since he preferred to avoid the indeterminacy of personal identity. Thus, we can understand how he came to his famous claim that personal identity is not what matters.

### 3. The Classic Narrative Approach: Origins

Throughout the 80's of the past century, the terms of the debate on personal identity within the OA began to be considered futile. Parfit's argument was incredibly forceful for those who accepted reductionism, but his conclusions were not so welcome. In this sense, it was inevitable that the focus of the debate shifted to a new approach that proved to be more productive. And this new approach was the CNA, which began to be developed after the publication of MacIntyre's *After Virtue* (2007). Nevertheless, we must consider that MacIntyre's work was not the cause of the rise of narrative approaches, but rather the contrary. Narrativism, as a broader trend within the humanities which interprets all human knowledge in narrative terms, gained strength in general in the early eighties (see Bruner, 1991:5). In this sense we should consider MacIntyre's theory more as a result of this trend than as a completely groundbreaking approach. Not in vain, Dennett published the first version of his theory of the narrative self in the same year as MacIntyre (Hofstadter & Dennett, 1981:348-52). In any case, as it happened with the OA, my interest does not lie in the different theories that were developed within the CNA. Instead, I will focus on the CNA at a higher level of abstraction, trying to stress its main theses and how they should be interpreted.

To begin with, there is a couple of things that we must consider regarding my reconstruction of the debate. The first one is that, according to my view, the CNA attempted to give an account of the problem of personal identity from a neo-Lockean perspective. Thus, in principle I will consider that CNA proponents were interested in holding the three claims stated in the introduction. This is not a trivial point, since contrary to what it may seem, not all of CNA proponents were specifically interested in this aspect. Some of them held these claims in a more or less explicit way (MacIntyre, 2007:217), and others even linked their positions to the previous debate (Schechtman, 1996:7-70). However, there were others who blithely accepted, for example, that personal identity is indeterminate (Dennett, 1991:422-3). In this sense my reconstruction of the debate will be a bit idealistic.

The second thing we must take into consideration is that an aspect common to most of CNA accounts was their rejection of reductionism and thus of claim (4). In this sense, I interpret that CNA's attempt to hold the three claims stated in the introduction departed from the intention to be true to our intuitions regarding the problem of personal identity, as they are stated in claims (2) and (3), without resorting to dualism, as Parfit had suggested.

Finally, an important feature of CNA is that most of its proponents did not only reject reductionism, but also the relevance of thought experiments to ques-

tions about personal identity. In this sense, most of them did not really put their accounts to the test, and so it is difficult to know if they succeeded in holding the three claims. Obviously, I take it that they did not, as I will show below. In any case it is impossible to know what some of them would have said about certain puzzle cases.

Having all this in mind, I will start my presentation of the CNA with two criticisms which its proponents made against reductionism. Then I will link each of these criticisms to the two main theses the CNA holds: those which Strawson labelled as the “*psychological Narrativity thesis*” and the “*ethical Narrativity thesis*” (see Strawson, 2008). The first criticism is based on the aforementioned objectivation of persons that reductionism implies, and which made OA proponents consider mental states as impersonal units that might be isolated from the rest of the mind to be removed, altered or transplanted from one mind to another without major complications. In response to this problem, the CNA proposes a more subjective approach to our consciousness. This approach takes into consideration that our experiences cannot be understood independently of the context in which they occur, since our brain organizes our mental life in narrative form. Thus, this response will take the CNA to the formulation of the psychological Narrativity thesis. The second criticism that I will consider concerns the fact that moral issues implied by the notion of personal identity never make sense out of the social context in which they arise. So, for example, questions about moral responsibility are meaningless when persons are removed from their environment. One cannot speak of moral responsibility if one does not take into consideration the other to which one must respond, and the only way to do so is through a narrative understanding of human life. This thought will crystallize in the ethical Narrativity thesis.

### 3.1. *The subjective dimension of experience and the psychological Narrativity thesis*

Perhaps the best way to understand the criticism regarding the loss that involves the objectivation of persons when giving an account of the problem of personal identity is the one Ricoeur presents when making his famous distinction between identity as sameness (*idem*) and identity as selfhood (*ipse*). That is, between identity understood as a response to the question “what am I?” and identity understood as a response to the question “who am I?” (Ricoeur, 1994:118). In proposing objective criteria of personal identity, OA supporters merely tried to find an answer to *what* we are, since the question about the “who” necessarily requires an approach to our existence that can account for our subjectivity. As we saw in the previous section,

all research on the topic of personal identity within the OA was done from the assumption that it was possible to make an *impersonal* description of persons. That is, that it was possible to say *who* we are after removing the “who”, since objectivity means ignoring, both in regard to our bodies and to our minds, the phenomenical quality of “being ours” that they have. The concept of quasi-memory, developed by Shoemaker and which is central to the OA (see note 2 above), is the best example of this depersonalization. The memory that belongs to no one cannot account for a “who” since it is precisely the “who” that we remove when *quasifying* it. To paraphrase Ricoeur, one’s own memories cannot be a particular case of quasi-memories since it seems incomprehensible how what is one’s own can be a particular case of the impersonal (Ricoeur, 1994:133).

According to CNA proponents, this was one of the reasons why the OA could not give a satisfactory solution to the problem, since what matters to us in personal identity is precisely the “who” question. “Who am I?” is a question directed to what subjectively characterizes me, what makes me different from others and makes me the person I am (Gutiérrez Aguilar, 2019). In this sense what matters to us is the identity that lies behind what we call an “identity crisis” (Taylor, 2001:27). The self only exists within a framework of subjective assessments and that is where one has to look for it. In Schechtman’s words, the problem with the OA was that it mixed what she calls the reidentification question and the characterization question, as questions about the logical conditions used to identify a person over time cannot provide answers to what matters to us when we wonder what makes us who we are (Schechtman, 1996:1-2).

Thus, the search for a way to get closer to our experience of the world that did not betray it by objectivizing it had to start from introspection, from an analysis of how we subjectively perceive reality. And the answer came almost simultaneously from two different sources: MacIntyre and Ricoeur said that we experience the world in narrative form, so if we consider our experiences as part of our narratives about our own lives, we will be taking into consideration the subjective perspective which had hitherto been ignored.

MacIntyre’s point is easy to understand and is based on the fact that to comprehend the behaviour of a person we must connect it with a variety of aspects that surround it, such as the reasons that led the person to perform that action and the context in which it takes place (MacIntyre, 2007:206). He explains it from a simple example. We can imagine a situation where a person is involved. To the question “what is he doing?” we can say truthfully: “digging”, “gardening”, “taking exercise”, “preparing for winter” or “pleasing his wife”. All these answers could be equally true and appropriate descriptions of the same situation, but

nevertheless to choose between these options and decide how we are going to describe it, we need to take into account the intentions of the person involved, because the action is not the same if the main reason that has led the man to the garden with a shovel is exercising or avoiding a quarrel with his wife, since they are enshrined in different contexts. Thus, we find it necessary to accept that every situation is immersed in a sequence in our mind, a narrative that connects it with all the aspects which are necessary for its understanding, and that without such a narrative the situation is unintelligible. It is therefore impossible to individuate mental states in the way that supporters of psychological criteria did when talking about memory implants and the like, but we have to consider complete narrative sequences.

For its part Ricoeur's thesis, which can be considered as complementary to that of MacIntyre, comes from his analysis of temporality. It reaches its greatest expression in the three volumes of his work *Time and Narrative*. There Ricoeur analyses the human experience of temporality and its relation to the activity of narrating, and he says that:

[...] between the activity of narrating a story and the temporal character of human experience there exists a correlation that is not merely accidental, but that presents a transcultural form of necessity. To put it in another way: *time becomes human to the extent that it is articulated through a narrative mode, and narrative attains its full meaning when it becomes a condition of temporal existence.* (Ricoeur, 1984, p. 52)

This is because, as MacIntyre said, the understanding of an action requires its disposition in a plot, which ultimately—and to avoid a problem of circularity in his theory of the three stages of mimesis (Ricoeur, 1984:71-6)—implies that there must be a pre-narrative structure of temporal experience (Ricoeur, 1984, pp. 59–60). That is, it is necessary to postulate an original temporal experience, shared by all human beings, which demands a narrative development (Ricoeur, 1984:74).

The narrative claim made by MacIntyre and Ricoeur was accepted willingly by cognitive psychologists, among which we can highlight Bruner, who in his *Actual Minds, Possible Worlds* references the French philosopher as he defends the existence of two primitive modes of thought: the paradigmatic—or logico-scientific—which is the one we use to understand the physical world and that leads to science; and the narrative—or intentional—which is the one we use to understand human behaviour (Bruner, 1986:11-43). Thus, according to Bruner, by explaining our lives in terms of desires, beliefs or intentions, we are necessarily situating ourselves as the characters in a narrative, and our identity is defined as we are the main character of a particular story: our autobiography (Bruner, 1987:15).

Throughout the eighties and until recently, due to both the influence of the work of these authors and to the peak of narrativism in general, the belief that persons necessarily understand their lives in narrative form has become a cliché that needs no further demonstration. In this sense, one of the main features of the CNA is its endorsement of what Strawson called the “psychological Narrativity thesis” (Strawson, 2008:189). Thus, we can say that CNA proponents hold a fifth claim besides the three claims stated in the introduction:

(5) Selves (or persons) subjectively experience and understand their lives in narrative form, so that their whole mental life, their consciousness, is unified through this narrative self-understanding.

### 3.2. *The social dimension of our existence and the ethical Narrativity thesis*

The second criticism that CNA supporters launched against the OA begins with the importance that the problem of personal identity has for moral issues. Indeed, it seems clear that in order to explain why and under what conditions a person is responsible for what he did in the past, despite any changes he may have experienced since then, we need to respond first to the question about the identity of that person, since one can only be responsible for one’s own actions. Thus, moral issues demand that we have a criterion of identity that considers the lives of persons as a unit with long-term existence while ensuring that their identity will always be determinate (see Schechtman, 1996:149). Since OA is not able to provide such a criterion, we should develop a different approach to the problem. Moreover, even if it were possible to develop a criterion of personal identity that would meet these requirements from the OA, which it is not, we must also consider that moral issues are meaningful only in the social space where they arise, so our criterion of identity must also consider that the lives of persons always happen within a community. As MacIntyre puts it, persons must always be able to respond to the imputation of strict identity because “I am forever whatever I have been at any time for others—and I may at any time be called upon to answer for it—no matter how changed I may be now” (MacIntyre, 2007:217). It is this social aspect of our existence that makes and demands our identity be determinate. Consequently, physical or psychological continuity will never be able to account for this fact since the unity of the life of persons does not depend solely on themselves. Following Ricoeur, only if we understand persons as characters in a narrative which is embedded in a social context, we can give them the required unity (Ricoeur, 1994:158).

But, once we have accepted the psychological Narrativity thesis, the problem we now have to face is not whether persons experience their life in narrative form, as this has been presupposed as a necessity, but which conditions should fulfil this narrative to shape a moral life. The first requirement, we have already said, is that the life of a person has to be gathered into a single narrative. The second requirement is that the narrative must somehow be open to accommodate the social dimension of persons. In fact, as we have seen, the first requirement is subject to the second, as the need for all our life to be gathered into a single narrative comes from the imputation of strict identity that our social existence demands. However, this has not been previously pointed out with appropriate strength, which has sometimes resulted in a misunderstanding of the CNA. In any case I will come back to this later. For now, I will focus on saying a little more about each of the two requirements as they were understood by the supporters of the CNA.

As to the first requirement, that our lives have to be gathered into a single narrative, it brings a problem, and it is that of the intelligibility of that narrative. It could be that our life has been so eventful that we are unable to tell an understandable narrative of it. This is precisely what happens when we say that our life is meaningless (MacIntyre, 2007:217) or that we have an identity crisis (Taylor, 2001:27). In such a case what happens is that we are unable to value who we are and what is important to us because our narrative is incomprehensible—that is, we lack an objective, a goal for our lives, which is a necessary condition to make a narrative intelligible. Thus, the way to give narrative unity to our existence must be to direct it toward a climax or *telos*. A road map which points out what we want to achieve will allow us to judge the situations in which we find ourselves and to act in the best possible way to achieve our goal. Only then our lives will acquire meaning, although obviously this meaning will only be understandable as the narrative's climax. And what better way to run our life than doing what is good for us, thus giving our existence the shape of a quest for the good (Taylor, 2001:51-2).

However, here the second requirement comes into play. Although the self is constituted in part by its own narrative self-interpretation, that interpretation is done by using a particular language. And since there cannot exist a private language, the idea of an isolated self makes no sense, for one is a self only among other selves constituted in the same linguistic context (Taylor, 2001:34-5). Thus, the language in which the selves are articulated implicitly carries the moral valuations of the society in which they exist, clearing the way for the introduction of the social dimension of persons in the CNA. This will be completed by MacIntyre, who moves from the linguistic level to the narrative.

MacIntyre points out that we become part of a narrative from the moment we are born. However, we do not do it in a scenario that we have chosen. Nor is the action in which we take part in of our authorship, as it is not in our own narrative in which we begin to exist, but in the narratives of others. We do not learn to speak until several months after birth, though it would be absurd to argue that we do not exist before that. Our parents give us a name and begin to tell stories about us, to interpret our behaviour and how it affects their lives from the moment we are born. From this perspective our own personal drama is constrained by and constrains other dramas of which we are also part, although our participation in them may be minimal. So, while we are the main character of our own narrative, we play only a supporting role in that of our parents, and just that of an extra in the narrative of the doctor who helped us to be born. However, although our role may be less important in others' narratives, we do not bow down to the script that others have written for us in their lives, but we can act on them of our own free will. Likewise, the other characters in our own drama write their lines themselves, so we must admit that we are never more than the co-authors of our own narrative (MacIntyre, 2007:213).

At this point it is important to note that since the psychological Narrativity thesis does not imply the need to have such a structured narrative of our own life, it is always possible that our narrative may not meet these requirements. We can recall with no surprise the main character of the novel by Max Frisch, *I'm Not Stiller* (2006), who after fleeing the country and attempting suicide refuses to accept both the name and the responsibilities of his previous life. In such a case, and since as we have seen our narrative is affected by the other narratives in which we participate, society may try to force that person to recognize his identity, as in the case of the main character of the novel, who is taken to court by the state of Switzerland and forced to accept his legal responsibilities. If this is the case, and as it happens in the novel the person involved finally recognizes its responsibility, we would be just talking about an unethical person who ought to restructure his life. However, it is possible that someone with a severe mental disorder refuses to accept the role that the rest of society imposes on him, or that some person is unable to offer a minimally structured narrative of his entire life. Considering these situations where a person is unable to provide a single, coherent narrative of his own life because of a mental illness, Schechtman says that although it is not necessary that all human beings have a narrative perspective on their own lives with the requirements we are discussing, it is necessary to achieve full personhood, so that ultimately not all human beings are persons (Schechtman, 1996:119).



Such considerations led CNA proponents to the endorsement of a sixth claim, which is equivalent to what Strawson called the “ethical Narrativity thesis” (Strawson, 2008:189):

(6) Narrative self-understanding is a necessary condition for moral existence, so that selves (or persons) are moral beings, and thus achieve full personhood, in virtue on their ability to understand their social lives in a narrative way.

However, we should note that there is a common misunderstanding regarding this claim. As I have tried to stress throughout this section, the main reason that led CNA theorists to hold this claim was the need to accommodate the social dimension of our existence, an aspect which the OA ignored. The claims made by MacIntyre, Taylor and Ricoeur regarding the importance of having a highly structured narrative of our own life oriented towards a defined goal were not an aim in themselves, but only a way to account for our social life.

#### 4. The Classic Narrative Approach: Crisis

Despite of its superiority with regard to the OA, in recent years there have been several criticisms against the CNA. Here I would like to focus precisely on those related to the way in which it accounts for the subjective and the social aspects of our existence. Specially, through Strawson’s criticisms in his “Against Narrativity”.

In this famous article, after identifying and describing both Narrativity theses, Strawson criticizes them harshly. To do so, he departs from a differentiation between two ways in which persons experience their being in time, which he takes to have a genetic basis, and which thus gives rise to two types of persons: Diachronics and Episodics (Strawson, 2008:190-1). Diachronics naturally figure themselves, when they consider them as selves, as entities that existed in the further past and will still exist in the further future, while Episodics, on the contrary, lack this sense. Although an Episodic can remember a past event and concede that it is an experience which was felt by him as a human being, he does not feel that such experience belongs to him as a self. He feels it alien to himself to the extent that he considers that it was experienced by a different self. This feeling could be interpreted as a failure of his self-perception, as Diachronics tend to do when faced with an Episodic. However, with the same authority we can assume that it is a reflection of a fact about what a self is, since we have no reason to doubt the reliability of his subjective feelings.

The question then is what the relationship between Diachronicity and Episodicity, as ways of experiencing our being in time, and Narrativity, as a way of understanding our life, is. And here Strawson distinguishes between two possible meanings of “narrative” and “Narrativity”. On the one hand, he admits that there is a sense in which certainly all of us understand ourselves in a narrative way. If we hold that making coffee involves narrative thinking because we have to think ahead, do things in the right order, and so on, then he is glad to admit that narrative thinking is a feature of every normal human being (Strawson, 2008, p. 198). However, he considers that this is a trivial claim, and that CNA cannot hold that what matters in survival and moral responsibility depends on this kind of narrative thinking. Even an Episodic can agree that it is possible to tell a narrative of his whole life as a human being, but that does not mean that he cares about his past existence as if it were *his* past. Consequently, if CNA proponents want to make a distinctive and non-trivial claim in this regard, they have to rely on a stronger sense of what “narrative” means. For example, they can hold that for something to be a narrative it must have the formal structure of classic stories which gives literary narratives the unity and coherence which trivial narratives lack (see Vice, 2003). If CNA holds that we understand ourselves in narrative form in this non-trivial sense, then they can claim that the narrative of our own life determines what matters to us in survival and moral responsibility. However, if they understand the term “narrative” in this non-trivial sense, then they have to admit that not all people understand themselves in narrative form. For example, Episodics do not have a narrative of their whole life which gives it the required unity and coherence. Neither do some Diachronics, who can accept that their past is theirs without positively grasping their life as a unity in a narrative sense (Strawson, 2008:200-1).

But although it is not true that everyone naturally experiences his own life as a narrative, it could still be argued that trying to do so should be taken as a moral duty, since according to the ethical Narrativity thesis that is the only way we can be moral beings and thus achieve full personhood. However, as Strawson notes, our moral qualities are not affected by our experience of time. According to Strawson, the idea that beats deep down in the ethical Narrativity thesis is that it is necessary to have a profound understanding of oneself in order to act in an ethical manner, and that this self-understanding can only occur in a narrative form. However, although we can concede that a certain degree of self-understanding is necessary to have a good life, it can happen without any narrative at all. One can delve into one’s own life without having an explicit link with one’s past, just as musicians can improve through practice sessions without recalling those sessions (Strawson, 2008:205). Similarly, one can be a good friend to someone regardless of the value

that one gives to past shared experiences, as friendship is shown in how one is in the present (Strawson, 2008:207). If we go even a little further, we can also see that it is not only unnecessary to have a narrative understanding of one's life to have a good life, but it can be harmful for some persons who feel that considering themselves as if they were characters in a narrative is to dehumanize themselves and to distort their own life (Strawson, 2008:205). Even if all the facts are right, these persons think that imposing on them a narrative structure is missing the truth, losing the right perspective in which their life should be understood. This sense of alienation and derangement can eventually cause serious emotional problems, so taking a narrative perspective, far from leading to an improvement in their life, can cause serious psychological harm to some persons.

Strawson's criticisms show that both Narrativity theses are false, at least in any non-trivial sense. And since they are the two claims of the CNA which distinguished it from other neo-Lockean accounts of the problem of personal identity, it is easy to understand why Strawson's article has had such an enormous influence in the debate. The core of his criticisms lie in the way CNA interprets the term "narrative". If they take it in a trivial sense, then the narrative through which we understand ourselves cannot bear the importance we attach to personal identity with regard to survival and moral responsibility, since it could not give our whole lives the required unity and coherence. Thus, although claim (5) could be true, claims (2) and (6) would be not, since our narrative self-understanding would have nothing to do with what matters to us in survival and moral responsibility. In of the contrary they interpret the term "narrative" in a non-trivial sense, then their claims would be too demanding, so that none of them would be true for some of us. Specially for Episodics, who do not understand themselves as selves (or persons) as entities with a long-term existence. Therefore, what Strawson criticisms show is that, in the trivial sense, CNA claims are either false or irrelevant. And that in the non-trivial sense, claim (5) is unable to account for the way some people subjectively understand themselves and claim (6) fails to account for the way our social relationships acquire meaning.

CNA theorists have been trying to offer a response to Strawson's criticisms for the last fifteen years. Some of them have succeeded to a certain extent. For example, Rudd (2007:63) noted that claim (5) could never be taken as trivial, even if we understand the term "narrative" in a trivial sense, since there are alternative conceptions of how human actions should be understood (for example, reductionism). He also stressed the fact that moral virtues imply certain kind of temporality which can only be assessed in a Diacronic way (Rudd, 2007:70). Thus, using Strawson's example, friendship could never be reduced to an Episodic kindness. However, de-

spite of its relevance, Rudd's defence leave unanswered the main point of Strawson's criticisms: if both Narrativity theses are true, then they have nothing to do with what matters in survival, since Episodics can accept that there is a trivial narrative about their whole life as human beings without being concerned about their pasts and futures. They would still feel that those moments were or will be experienced by a different self.

## 5. The New Narrative Approach

The best response to Strawson's criticisms was offered by Schechtman. However, as we have seen, his criticisms were directed towards the very basis of the CNA. Thus, Schechtman's response is not a simple revision of the CNA, but it implies a full reassessment of the problem of personal identity from a neo-Lockean perspective. In this sense I consider her response as the starting point of the NNA.

As Schechtman was well aware of, what Strawson's criticisms reveal is that there is an underlying tension in the CNA between the subjective and social aspects of our existence. CNA theorists wanted to pay attention to the way we subjectively experience our own lives, and so they posited that we understand ourselves in a narrative way. At the same time, they wanted to account for the complexities of our social existence, so they argued that the narrative by which we understand ourselves must give a strong unity and coherence to our life, so that it can respond to the imputation of strict identity. But while there are, according to Strawson, some people who subjectively experience themselves as entities which existed in the further past and will be there in the further future—Diachronics—there are others who experience themselves in a somehow less extended way—Episodics. That being the case, if we posit a single entity constituted through a single narrative, for Episodics this narrative will be either too short to account for the social dimension of their existence or too extended to be true to the way they subjectively experience their life (Schechtman, 2007:167-9).

Schechtman's proposal to solve this tension consisted in establishing a distinction between two kinds of narratives through which we understand ourselves and which thus give rise to two kinds of entities that we are besides human beings: selves, which would account for the subjective dimension of our lives; and persons, which would do the same regarding the social aspects of our existence. Within this new framework, selves would be conceived as the subject of experiences, an inner mental entity with which we strongly identify, so that an experience belongs to ourselves as a self only if we identify and care about or take interest in it. That is, if one

has “empathic access” to it (see Schechtman, 2001). For their part, persons would be conceived as the bearer of certain social capacities (Schechtman, 2007:169). So, for her, to be a person one must recognize oneself as an entity which has a long-term existence. One must acknowledge that one’s past actions have implications for one’s current rights and responsibilities and that one’s future will also be impacted by one’s present actions. In any case this does not mean that one must identify oneself in a deep sense with his past or his future experiences. But one has to recognize them as relevant for one’s present. As Schechtman puts it, “I need not identify with the self who decided to buy the sports car, but if I signed the loan, I need to recognize that it is mine to pay, and that my credit will be impacted if I do not” (2007:170).

This kind of response to Strawson’s criticisms has been highly welcome among philosophers interested in the problem of personal identity (see e.g., Johnston, 2010; Menary, 2008; Stokes, 2013; Zahavi, 2007). Thus, I take that the distinction between selves and persons is the main feature of the NNA, so that it is characterized by its endorsement of a modified version of claim (1):

(7) We have to distinguish between at least three kind of entities that we are, mental selves, social persons, and biological human beings.

However, despite the acceptance of this claim, there are still huge disagreements about how this distinction between selves and persons should be understood, and what consequences it carries for claims (2) and (3). That is, about our beliefs that personal identity is determinate and that it is what matters in survival and moral responsibility, since it is not clear which of the three entities that we are according to the NNA must respond to the imputation of strict identity and which should matter to us in survival and moral responsibility. As this is a historical account, this article must end with an open position in this regard. However, I cannot fail to mention what seems to me as the main agreement within the NNA.

As we have already said, the distinction between selves and persons has as its main objective to account for the subjective and social dimensions of our existence. Thus, given that moral issues only have sense within the social space in which they arise, most of NNA proponents agree that persons are the entities that should matter to us when we care for moral responsibility. In this sense, it also seems that NNA proponents are willing to accept MacIntyre’s point with regard to the role that the narratives told by others about ourselves play in the constitution of our own personal identity, which would also let NNA proponents defend that the identity of persons is determinate (see Muñoz-Corcuera 2021a). However, this does

not mean that NNA proponents hold that the identity of selves is indeterminate, or that selves do not matter when we care for moral responsibility. It does neither mean that they agree that persons are the entities which matter to us in survival. These questions are still open and should be taken as the main duties for the development of the NNA.

## BIBLIOGRAPHY

- Bruner, J. S. (1986). *Actual Minds, Possible Worlds*. Harvard University Press. DOI: <https://doi.org/10.4159/9780674029019>
- Bruner, J. S. (1987). Life as Narrative. *Social Research*, 54(1), 11-32.
- Bruner, J. S. (1991). The Narrative Construction of Reality. *Critical Inquiry*, 18(1), 1-21. DOI: <https://doi.org/10.1086/448619>
- Butler, J. (2008). Of Personal Identity. In J. Perry (Ed.), *Personal Identity* (pp. 99-106). University of California Press.
- Dennett, D. C. (1991). *Consciousness Explained*. Brown and Company.
- Flew, A. (1951). Locke and the Problem of Personal Identity. *Philosophy*, 26(96), 53-68. DOI: <https://doi.org/10.1017/S0031819100019203>
- Frisch, M. (2006). *I am not Stiller*. Dalkey.
- Grice, H. P. (1941). Personal Identity. *Mind*, 50(200), 330-350. DOI: <https://doi.org/10.1093/mind/L.200.330>
- Gutiérrez Aguilar, R. (2019). Facing Boredom: Essential Indexicals and Narratives of the Self. En J. Ros Velasco (Ed.). *Boredom Is in Your Mind: A Shared Psychological-Philosophical Approach* (pp. 123-134). Cham-Switzerland: Springer Nature. DOI: [https://doi.org/10.1007/978-3-030-26395-9\\_8](https://doi.org/10.1007/978-3-030-26395-9_8)
- Hofstadter, D. R., & Dennett, D. C. (Eds.). (1981). *The Mind's I: Fantasies and Reflections on Self and Soul*. Penguin.
- Johnston, M. (2010). *Surviving Death*. Princeton University Press. DOI: <https://doi.org/10.1515/9781400834600>
- Locke, J. (1975). *An Essay Concerning Human Understanding*. Oxford University Press.
- MacIntyre, A. (2007). *After Virtue: A Study in Moral Theory (Third Edition)*. University of Notre Dame Press.
- McDowell, J. (1997). Reductionism and the First Person. In J. Dancy (Ed.), *Reading Parfit* (pp. 230-250). Blackwell.
- Menary, R. (2008). Embodied Narratives. *Journal of Consciousness Studies*, 15(6), 63-84.

- Muñoz Corcuera, A. (2009). El caso de Memento: Una memoria nietzscheana en el cine postmoderno. *Bajo Palabra: Revista de Filosofía*, 4, 181-190.
- Muñoz-Corcuera, A. (2021a). Persistence Narrativism and the Determinacy of Personal Identity. *Philosophia*, 49(2): 723-739. DOI: <https://doi.org/10.1007/s11406-020-00265-8>
- Muñoz-Corcuera, A. (2021b). Narrativism, Reductionism, and Four-Dimensionalism. *Agora: Papeles de Filosofía*, 40(2), 63-86. DOI: <https://doi.org/10.15304/ag.40.2.6713>
- Nozick, R. (1981). *Philosophical Explanations*. Harvard University Press.
- Olson, E. T. (1997). *The Human Animal: Personal Identity Without Psychology*. Oxford University Press.
- Parfit, D. (1984). *Reasons and Persons*. Oxford University Press.
- Reid, T. (2008). Of Mr. Locke's Account of Our Personal Identity. In J. Perry (Ed.), *Personal Identity* (pp. 113-118). University of California Press.
- Ricoeur, P. (1984). *Time and Narrative I*. University of Chicago Press. DOI: <https://doi.org/10.7208/chicago/9780226713519.001.0001>
- Ricoeur, P. (1994). *Oneself as Another*. University of Chicago Press.
- Rudd, A. (2007). In Defence of Narrative. *European Journal of Philosophy*, 17(1), 60-75. DOI: <https://doi.org/10.1111/j.1468-0378.2007.00272.x>
- Schechtman, M. (1996). *The Constitution of Selves*. Cornell University Press.
- Schechtman, M. (2001). Empathic Access: The Missing Ingredient in Personal Identity. *Philosophical Explorations*, 4(2), 95-111. DOI: <https://doi.org/10.1080/10002001058538710>
- Schechtman, M. (2007). Stories, Lives and Basic Survival: A Refinement and Defense of the Narrative View. In D. D. Hutto (Ed.), *Narrative and Understanding Persons* (pp. 155-178). Cambridge University Press. DOI: <https://doi.org/10.1017/CBO9780511627903.009>
- Shoemaker, S. (1970). Persons and Their Pasts. *American Philosophical Quarterly*, 7(4), 269-285.
- Shoemaker, S. (1997). Parfit on Identity. In J. Dancy (Ed.), *Reading Parfit* (pp. 135-148). Blackwell.



Stokes, P. (2013). Will it be me? Identity, concern and perspective. *Canadian Journal of Philosophy*, 43(2), 206-226. DOI: <https://doi.org/10.1080/00455091.2013.826054>

Strawson, G. (2008). Against Narrativity. In *Real Materialism and Other Essays* (pp. 189-207). Oxford University Press. DOI: <https://doi.org/10.1093/acprof:oso/9780199267422.003.0008>

Taylor, C. (2001). *Sources of the Self: The Making of Modern Identity*. Harvard University Press.

Vice, S. (2003). Literature and the Narrative Self. *Philosophy*, 78(303), 93-108. DOI: <https://doi.org/10.1017/S0031819103000068>

Williams, B. (1957). Personal Identity and Individuation. *Proceedings of the Aristotelian Society*, 57, 229-252. DOI: <https://doi.org/10.1093/aristotelian/57.1.229>

Williams, B. (1970). The Self and the Future. *The Philosophical Review*, 79(2), 161-180. DOI: <https://doi.org/10.2307/2183946>

Zahavi, D. (2007). Self and Other: The Limits of Narrative Understanding. In D. D. Hutto (Ed.), *Narrative and Understanding Persons* (pp. 179-201) Cambridge University Press. DOI: <https://doi.org/10.1017/S1358246100009668>

---

DOI: <https://doi.org/10.15366/bp2021.28.001>

Bajo Palabra. II Época. N°28. Pgs: 21-46

