BALKAN JOURNAL
OF PHILOSOPHY

Vol. 13, Issue 1, 2021

The issue is sponsored by the Bulgarian National Science Fund, competition “Bulgarian Scientific Periodicals – 2020”, project No: КП-06-НП2/11 from 02 Dec. 2020

TABLE OF CONTENTS

Discussion

Book reviews

SPECIAL ISSUE OF THE BALKAN JOURNAL OF PHILOSOPHY

“THE USES OF NARRATIVES: FROM PRACTICE TO EPISTEMOLOGY AND BACK”

The uses of narratives, the story-like accounts of particular subjects, are well recognized in various areas of human knowledge and practice. Narratives are ubiquitous both in daily life and in sophisticated professions such as medicine and criminal investigation. They are also common in art, science, and the humanities. The ubiquity of narratives has been associated with the important roles they play: from yielding understanding to serving as justifications for certain actions or as arguments in favor of chosen theories or stances. The multiple uses of narratives, however, raise questions which cannot be easily settled and which, therefore, have become a subject of continuous debate. Do narratives deliver truth or are they intrinsically “just-so stories”? Related to this epistemological question is the methodological one regarding the status of so-called narrative-based medicine: should it be treated as a strict alternative to the currently dominant evidence-based medicine or should it be viewed as its important complement? While it has been broadly accepted that narratives provide understanding, it remains controversial whether they do so because they offer a kind of genuine explanation. Should narratives have to meet certain criteria in order to claim an explanatory role? Or, if instead we take the stance that narratives provide non-explanatory understanding, in virtue of what are they able to do so ? It seems to be beyond question that narratives can play a justificatory role. It is uncertain , however, whether narrative arguments are indispensable to scientific discovery or have been introduced into scientific discourses simply to increase confidence in the established theories that these narratives seek to defend.

Given the salience of the epistemological and methodological controversies surrounding the uses of narratives, it is surprising that so far so little effort has been devoted to addressing these controversies in a systematic way. The aim of the special issue of Balkan Journal of Philosophy for 2021 is thus to contribute to the developing investigations into the epistemology of narratives and their practical applications. We invite papers that address epistemological and methodological questions concerning the uses of narratives, both from a general theoretical perspective and in the context of specific areas of science, art, the humanities, and various professional practices.

From the Editorial Board of
Balkan Journal of Philosophy

Articles

Invited paper

ON THE COMPATIBILITY OF COMPETING NARRATIVES

David Weberman
Department of Philosophy, CEU, Vienna
e-mail: WebermanD@ceu.edu

Abstract:
We reconstruct past events, whether in history or in everyday life, in the form of narratives. Yet narratives describing one and the same set of events can and do differ. What is the relation between these different narratives? Must they necessarily conflict? When are they compatible and when not? If we can tell stories differently without getting the facts wrong, what constraints can there be for judging the adequacy of competing narratives?

Key words: narrative, truth, divergence, complementarity, incompatibility.

Narratives can be divided into two kinds: fictional and non-fictional. My paper concerns non-fictional narratives. We get non-fictional narratives not only from professional historians but from ordinary people relaying stories about themselves and the events that occur in their everyday lives. Narrative is pervasive. It is probably the primary form used to “capture” the past. By capturing the past through narrative, we come to understand much about the present: Stories track the progression of choices, experiences, and interactions that have culminated in the present, allowing us to access an ordered conception of events that are no longer observable. Thus, narrative is at the heart of how we make sense of the world and our unique place within it. Despite the important role we assign to narrative, it remains unclear whether we can in fact offer a satisfactory account of the conditions under which competing narratives are true or false and adequate or inadequate. The lack of clarity and consensus on this issue leads to a number of questions that I intend to address in the following paper. First, when are divergent narratives mutually compatible and when are they incompatible? Second, when competing narratives are judged to be incompatible, under what conditions might one narrative be regarded as epistemically reliable and the other as epistemically unreliable? In answering these questions, I hope to provide some initial insights that shed light on the adequacy of narrative as a system of meaning-making and knowledge acquisition.

The stories we tell about past events typically diverge from one another, sometimes more, sometimes less. This certainly occurs in the world of historical scholarship. It also occurs in everyday life, when two or more parties tell stories about the same subject matter or events. How should we assess such divergent narratives? To be sure, divergent narratives often contradict one another at the level of individual facts. If one narrative maintains that two gunshots were fired and another narrative maintains there were three, then at least one of the two narratives contains a false assertion and is epistemically defective, whereas the other may or may not be. But consider now the possibility that two divergent narratives covering the same events or the same distinct segment of time and space do not contradict one another at the level of individual facts. What is the relation between two such divergent narratives? Are they compatible or not? Are they always on equal epistemic footing?

It can occur that two divergent narratives covering the same events are thoroughly compatible. They do not contradict one another on simple facts. They are not only compatible (logically consistent with one another), but complementary (they add to and fill each other out). If they are complementary, they give us more information about, and a fuller picture of, their subject matter.

One might then think that the relation between divergent narratives of the same subject matter is either a) contradictory at the level of individual facts or b) mutually compatible and complementary. Yet one should note, and it certainly has been noted, that the relation between divergent narratives is more complicated than this. Philosophers of history, such as Arthur Danto and Louis Mink, have made the point that historical narratives are different from chronicles. They do not only list events, they attempt to shed light on the way these events hang together. In doing so, divergent narratives can have a more complicated relation. They can compete, epistemically, at what one might call the narrativistic level.

This points to the primary question I will be exploring, namely: What is the nature of narrativistic truth? Here, I understand ‘narrativistic truth’ to mean truth not at the level of individual truth-apt assertions, but at the level of narrative, that is, the holistic assemblage of individual assertions into a broader system of meaning. I argue that the standards according to which divergent narratives either complement or rival one another vary depending on the type of non-fictional narrative that is put forward. I distinguish between four types of narrative and argue that the last type, what I will call a ‘narrative that functions to shape an identity,’ or, for short, an ‘identity-creating narrative,’ has a different nature than the other three types – one that gives it a certain hybrid status.

I. Compatibility, narratives and perspectives

Before exploring rival divergent narratives, I want to say something briefly about complementary divergent narratives. It seems to me that when divergent narratives are complementary, they bear a resemblance to divergent perspectives of one and the same thing. In the case of perspective – here I have in mind visual perspective – how something looks covaries with the location of the perceiver. In the case of narrative, how something was – i.e., what happened – covaries with certain features of the narrator, such as but not limited to the narrator’s experience, knowledge, or focus of attention.

The independent variable, whether in visual perception or in narration, underwrites a certain plurality of results that need not contradict one another, but may happily complement one another. As David Wiggins puts it:

> “Perspective is not a form of illusion, distortion, or delusion. All the different perspectives of a single array of objects are perfectly consistent with one another. Given a set of perspectives, we can recover, if only they be reliably corrected, a unified true account of the shape, spatial relations, and relative dimensions of the objects in the array.”[1]

The same holds true for narrative. Divergent narratives can enrich one another and lead to a fuller account of past events. But they don’t always do so.

An illustration of the complexity of comparing narratives is provided by Akira Kurosawa’s much cited 1950 film, *Rashomon*. Although it is fictional, we might consider it for a moment as if it were factual. It tells the story of a chance meeting in the woods between a bandit and a samurai and his wife which ends with the murder of the samurai. Divergent stories are told by the eye-witnesses and others regarding what they saw, or believe they saw, and what happened. What is interesting is that although the narratives do differ significantly at the level of individual facts, they include different events, aspects of events, and readings of aspects of events, some of which are compatible with one another; to the extent that they are compatible, it appears these divergences might contribute to a fuller story of what actually happened. It seems to me that one would need a detailed scorecard to begin to sort this out, to see which items complement and which items contradict

[1] David Wiggins, “Truth, Invention and the Meaning of Life” in Proceedings of the British Academy (1977), p. 350–351.

one another. In fact, this is also what courtroom judges do with many disputes and what we all do when trying to settle things in our daily affairs. For my purposes here, I want to isolate a specific issue, so I will consider divergent narratives only to the extent that they do not conflict at the level of individual facts. Given these conditions, I will examine in what sense such narratives can be said to stand in rivalry rather than in complementarity with one another, and interrogate what the nature of that rivalry is.

II. Narrativistic rivalry: Explaining or understanding why

Divergent narratives set in the same time and place do not stand in rivalry with one another when they simply pay attention to different things. Consider two narratives describing interwar Germany; one narrative focuses on its politics and the other on its cultural developments, yet despite substantive differences, they are entirely compatible and complementary. So when do they conflict at the narrativistic level? The answer comes as no great surprise. They are in disagreement when they aim at *explaining* the same thing, (e.g., explaining why something happened or why a certain outcome arose) but offer importantly different accounts. For example, after a car crash, a policeman will ask each of the two drivers involved for an account of what happened. The two accounts may make use of the exact same facts and both drivers may utter only true assertions; however, insofar as the two divergent narratives try to explain *why* the crash happened, their divergence may amount to rivalry. This is because each driver may perceive the same discrete facts and interpret the relationships between these facts in importantly different ways, and may thus offer very different narrative explanations for why the crash occurred. The same goes for divergent narratives about why a marriage soured or why it thrived. At some level, such narratives aim at providing the best explanation possible.

A few points need to be made here.

First, narratives are not always explicit about what they seek to explain. Furthermore, complex narratives usually explain more than one thing and may be composed of various parts adding up to a whole that lends itself to several different interpretations. As such, exploring the nature of rivalry between two narratives that are compatible at the level of bare facts will depend on identifying or discovering the explicit and implicit explanatory goals of each narrative.

Second, explaining why something happened or why a certain outcome resulted need not be construed as identifying tight causal connections under some sort of covering-law model of explanation. Explaining why might only consist in specifying the conditions that made a certain event or outcome probable or in showing how an event hangs together with the events that preceded it. In other words, explaining why an event, E, occurred might be a matter of selecting and expounding the facts that provide some minimum *understanding* of how E came to be.

Third, when divergent narratives explain the same thing differently, but do not contradict one another at the level of individual facts, the relation that they stand in is not one of a stated logical contradiction in the strict sense of the term, but one of narrativistic disagreement – incompatibility or rivalry. Such rivalry makes it natural to think that one narrative is epistemically superior to the other. In fact, I think it would be unnatural or at least puzzling to insist that although they explain the same thing differently, they are on equal epistemic footing. Explaining or understanding why is a competitive business.

Fourth, presumably they explain the same thing differently because they include different (although not necessarily incompatible) events in their narratives or describe the same events in importantly different ways.

Divergent narratives are not always rivals. As mentioned, they are not rivals when they discuss and explain different events or aspects of events that happen to occur at the same time and place, as with the earlier example of interwar Germany. Furthermore, not all narratives are explicitly or even implicitly explanatory. For example, when asked what I did today, I might tell a story that consists of little more than: First, I went out for a coffee; then I took a walk, did some shopping, read the newspaper, and took a nap. Let me call such a narrative an ‘innocent narrative.’ It has no explanatory intent. It is coherent as long as its key elements, its subject matter and setting (place and time), pose no incongruities. If there are divergent narratives in such cases, and they contain no false assertions, then as far as I can see they are always complementary and never rivalrous.

III. Two other types of non-innocent narratives

So far we have discussed two broad categories of non-fictional narratives: Innocent narratives, which at the narrativistic level are always compatible, and non-innocent narratives in the form of explanatory narratives that are always or at least typically in a relationship of rivalry. In addition, there are other types of non-innocent, non-fictional narratives besides those that aim at explaining or understanding why. I can think of two further non-innocent types.

One is moral. I have in mind a narrative of past events that aims to decide a moral question about an outcome produced by those events. It does not aim to show why some event or outcome occurred but whether that outcome is characterized by a certain moral praiseworthiness or blameworthiness. I won't go into this in any detail because I want to focus on the issue of compatibility and rivalry. Yet I would suggest that morally-motivated narratives are like explanatory narratives. They are rivalrous rather than complementary because one or the other divergent narrative can be criticized as less suited to the task of supporting a moral judgement concerning events described. Divergent moral judgements can be evaluated in this way because they serve a purpose other than enumerating events: They do not necessarily explain why a certain outcome has materialized, but they do provide evidence that allows us to consider whether this outcome is morally right or wrong. Unlike innocent narratives, they have a test to pass beyond that of not advancing any false assertions.

Another type of non-innocent narrative is practical. It tells a narrative with the purpose of recommending a certain future course of action. An explanatory narrative is backward-looking because it tells us why something happened. A moral narrative might be seen as present-oriented in that it provides information relevant to making judgments about present moral standing. A practical narrative is future-looking in that it fosters deliberation about what we ought to do next. Like other non-innocent narratives, it faces a test beyond getting the individual facts right. Since its purpose is to provide a narrative that gives the best account for deciding which future action is most prudent, it must be judged in reference to this standard.

To sum up: While innocent narratives can only disagree at the factual level, non-innocent narratives can also disagree at the narrativistic level, but, of course, there must be something to disagree about. I have suggested that there are at least three things to disagree about: why something occurred; what the moral status of a certain outcome is (e.g. whether the outcome in question is right or wrong); and finally, what the most suitable course of future action consists in. When non-innocent narratives diverge, we are forced to decide which narrative is superior or inferior, and therefore,which one ought to be criticized and why. In such cases, the constraints that allow us to determine what counts as a good or adequate narrative may be either explanatory, moral, or prudential.

IV. Narratives of Identity

In the last part of this paper, I would like to move on to yet another type of non-fictional narrative, namely, identity-creating or identity-forming or identity-bolstering narratives – in short, identity narratives. What we take our identity, or rather our identities, to be (i.e. who we take ourselves to be) does not just consist in some ordered list of our present features. Rather, identity is bound up with stories we tell about our past. I think it is fair to say that identity, both individual and collective, is always built up from narrative. Identity-narratives are not innocent because they serve a (often contentious) purpose beyond simply retelling events in a certain fragment of time and place. They are not haphazard, but may conform to a logical temporal order or be variously structured according to accidents of attention and memory. In either case, they are organized around, and serve the purpose of, expressing what we take ourselves to be.

Are identity narratives subsumable under the other three types of non-fictional, non-innocent narratives? Although identity narratives are never altogether detached from one or more of the other three types of non-innocent narratives (for reasons I'll return to shortly), it does not appear that they are reducible to any one such type or a combination thereof. Indeed, I argue that they introduce a new element which makes them a very special type of narrative. Who we are, or who we take ourselves to be, is not completely constrained by anything outside of our own choices. In other words, there is a voluntaristic element to identity narratives that is absent from other narratives. Who I am is, to some extent, up to me. And the story I tell, which makes me who I am, and how I tell that story, are also to some extent up to me. So, identity narratives are to some extent not constrained by anything outside of my own choices. For example, it seems that whether I identify with my nation of origin or whether I identify with a certain sexual or religious identity is a matter that is up to me. Similarly, whether Turkey chooses to identify as European or non-European might well be constrained by nothing outside of its own collective will – at least to some extent.

Now, I say "to some extent" because identity narratives are not unrelated to the other non-innocent narratives. Who I am is, in particular, a function of the unique stories I can tell about how I have become who I am, and the events detailed in these stories are in fact tethered to reality. Thus, an identity narrative can be criticized if it gets the facts wrong. But, insofar as it leans on explanatory, moral, and practical narratives, an identity narrative can also be criticized if it gets things wrong on a narrativistic level. Since identity narratives do lean on and connect with these other narratives, criticism at this level may be justified.

Given all of this, the question of who I am typically bears some connection not only to individual will (i.e. to what I will myself to be) but also to the way we apprehend lived experience (i.e. what has happened in my life and how I understand these events to be connected to what I have become). Consider as an example: If the Spanish take themselves to be Christian on the grounds that the Islamic presence in Spain was a mere disruption which made no positive contribution to Spanish history and culture (as some earlier Spanish history books contend), then their identity is not merely a matter of who they will themselves to be but also of how they understand and explain their present culture. Their narrative may be flawed at the level of bare facts, since they assert falsehoods, but also at the narrativistic level, because the connections they draw based on these falsehoods create a narrative that does not correctly *explain* how their culture really did develop. The same is true of an America identity based on a narrative that omits the history of the conquest of land through slaughter and the history of slavery. Such an identity rests on a narrative whose explanation of who we are today is narrativistically false.

It seems that identity is also tied to moral and practical narratives as well. For example, Palestinian identity is tied to a narrative centrally based on recounting what has happened and understanding how these events explain the present, but it is equally tied to a narrative about the Naqba (the disaster), which is concerned with showing that a moral wrong has occurred. Identity narratives may also be shaped by practical questions regarding the right course of action to pursue going into the future.

All of this treads on notoriously difficult terrain. For example, can a white person growing up in a largely white milieu self-identify as Black merely as a matter of will, and what kind of narrative could be told to underwrite that identity?

I don't have answers to such specific questions. What I've tried to do here is to provide a basic schema for thinking through the following questions regarding narrative truth: 1) When are narratives compatible and complementary and when are they rivalrous? 2) What are the different types of rivalrous narrative (explanatory, moral, practical) in virtue of which some narratives can be preferred over others? 3) What makes identity narratives both a hybrid form of the preceding narratives and characteristically different from any one of them or any combination of them (i.e. what is the nature of the voluntaristic aspect of identity narratives)?

Finally, I would like to suggest that identity narratives are highly controversial for this very reason. They have a dimension according to which they can be criticized for not conforming to the constraints governing adequate explanatory, moral, and practical narratives. Yet, they have a voluntaristic dimension which eludes such constraints. Much depends, then, on recognizing the difference between what is not up to us (explanatory truth, moral truth, practical know-how) and what is up to us (our self-determining identities that allow us to range over the narratives we choose to tell about ourselves). Thus, in navigating the constraints and freedoms of constructing and understanding the self through identity narratives, one of two mistakes might be made: either narrativistically misrepresenting the past or else underestimating one's own powers to determine the future.

References

Danto A (1985) *Narration and Knowledge*. New York: Columbia University Press.

Mink L O (1987) Narrative Form as a Cognitive Instrument. In: *Historical Understanding*, ed. Brian

Fay, Eugene O. Golob, and Richard T. Vann. Ithaca, 183–203. New York: Cornell University Press.

van den Akker C (January 2013) Mink's Riddle of Narrative Truth. *Journal of the Philosophy of History* 7(3): 346–370.

SELF AS ONE AND MANY NARRATIVES

Joseph Ulatowski
University of Waikato, New Zealand
e-mail: julatows@waikato.ac.nz

Abstract:
There are different approaches to the narrative self. I limit myself to one approach that argues narratives have an important role to play in our lives without it being true that a narrative constitutes and creates the self. My own position is broadly sympathetic with that view, but my interest lies with the question of whether there is truth in the claim that to create one's self-narrative is to create oneself. I argue that a self-narrative may be multiply realised by the inner self—impressions and emotions—and the outer self—roles in work and life. I take an optimistic attitude to the idea that narrative provides a metaphor that may stimulate insight into the nature of self if we accept a plurality of narrative selves. This paper mines a vein of research on narratives for insights into selves without being bewitched into accepting implausible conclusions.

Key words: self, narrative self, pluralism, practical reason.

"I look quiet and consistent, but few know how many women there are in me"
— Anais Nin

1. Introduction

The concept of self may be put to different uses, depending upon the philosophical area of interest in which it functions. For example, in questions of personal identity through time, we may ask: *If we were able to split one's brain in half and implant one-half in another body, would the self be split between the two bodies or would the self be retained in the original body?* A philosopher may be compelled to respond that the original self is lost or, following the work of Derek Parfit (1984), that this shows that the question over one's self does not matter. In biomedical ethics, we may ask: *If someone was seriously injured and unable to regain consciousness or to be able to live something like the life she was living before an accident, is the self before the accident retained afterwards?* When we are presented with such issues, we begin to wonder what a self is and we become quite sensitive to the details of what is constitutive of a self.

Discussion of what a self is began in earnest with discussion of self-consciousness. Consider Descartes' assertions in the second of the *Meditations* that, "*I am, I exist*, is necessarily true whenever it is put forward by me or conceived in my mind" (Descartes 1641/1998, 80), as well as in his *Discourse* and *Principles*: "I think, therefore I am" or "*cogito ergo sum*" (Descartes 1637/1998, 36; Descartes 1644/1998, 162, respectively). For Descartes, the *cogito* is something of which we can be absolutely certain. It is indubitable and can be known non-inferentially, e.g., intuitively. Locke accepts this intuition-based account when he claims, "*we have an intuitive Knowledge of our own Existence*, and an internal infallible Perception that we are. In every act of Sensation, Reasoning, or Thinking, we are conscious to ourselves of our own Being" (Locke 1700/1975, IV.ix.3). Descartes and Locke can be interpreted as accepting the view that there is an inner perception of the self or "inner self," but Hume notably disagrees:

> [T]here are some philosophers, who imagine we are every moment intimately conscious of what we call our self […] For my part when I enter most intimately into what I call *myself*, I always stumble on some particular perception or other, of heat or cold, light or shade, love or hatred, pain or pleasure. I never can catch *myself* at any time without a perception, and never can observe anything but the perception. (Hume 1739-40/1973, I.iv.§6)

Whereas Descartes and Locke were firm believers in an inner self, Hume was not, dismissing it because there is no impression, or perception, of oneself.

My contention is that we need to give full recognition to the subtle interplay between inner and outer in the production of a model of a narrative self. Such an interplay could be considered to fall somewhere between Descartes and Locke on the one hand and Hume on the other. In particular, we should avoid jumping to the conclusion that the inner self is the true self and the outer self is merely a reasonable facsimile of an inner self, maybe a behaviouristically interpreted self specially crafted by someone else. Just as wrong is the view that the outer self is the true self. Surely, the accurate picture is likely to be a composite one. This is a theme that will develop in the context of an inquiry into approaches to the self using the concept of a narrative, an approach that some might think would emphasise the inner self.

The notion of a narrative is currently widely employed in explanations and descriptions in a wide variety of domains, and in taking a narrative approach to the self there are a variety of options. In this paper, I narrow my focus to one family of approaches. Peter Goldie (2012) is persuasive in arguing for a certain kind of middle position concerning narratives and persons, namely that narratives have important roles in our lives without it being true that a narrative constitutes and thus creates the self, or that a narrative will determine personal identity in the sense of setting the identity conditions of persons through time. My own position is broadly sympathetic with that view, but my interest lies with the question of whether there is some element of truth in the claim that to create one's self-narrative *is* to create oneself. I want to take an optimistic attitude to the idea that narrative provides a metaphor that may stimulate insight into the nature of self. I should like to mine the vein of research on self-narratives for insights into persons and selves without being bewitched into accepting implausible conclusions.

2. The Inner and the Outer Self

Common in discussions of the self is the distinction between an inner and outer aspect. We typically speak of an inner life as composed of a person's observations, thoughts, and feelings. The inner life tells us of how things seem to be for that particular person. The inner life is private and accessible only to the person whose experiences they are. There are no means of sharing experiences of our inner life with others except through the expressive nature of language. Still, no matter how expressive we may be with our oral communication about our inner experience, something is missing from such illustrations. It is reasonable to assume that there is a privileged perspective that is ours and ours alone.

By contrast, there is the outer life of a self that consists of one's action or behaviour. The outer life is publicly available for others to see. The people around us observe what we do, and we observe what others do. Of course, one person's outer life is going to be represented by those others and thus appear in their own inner lives, modulated by their own interests and perspectives.

Such a distinction between the inner and outer self may be located in a broad variety of philosophical, scientific, and religious discourses, and indeed, we encounter it regularly in our everyday experience of life. Where we were feeling tentative and hesitant, we may learn that others saw us as confident and decisive, or perhaps it was the other way around. Or, we felt we were being witty and amusing when we learn later from others that they found us laboured and tedious.

Different characteristics of persons behave differently in relation to the distinction between the inner and outer self. Some characteristics appear to be defined in terms of their effects on others. Thus, the characteristic of being amusing can be defined as that of creating amusement in others. If others fail to be amused in spite of our best effort to be so, then, alas, we are not amusing precisely because of our failure to create the feeling of amusement in others. Our inability to have an effect on others suggests that outward manifestations of the inner self sometimes fail.

Even so, paired with the notion that is plausibly defined in terms of external characteristics there is a corresponding internal notion. A person may have a sense of humour (wicked perhaps) and that describes how their thoughts and utterances originate internally even though, not having the right audience, the humour is not appreciated. To take an example that initially goes in the other direction, we might want to say being anxious is definable simply in terms of an inner feeling characteristic of the inner self. We can appear to be calm and behave in ways that are consistent with being confident while feeling anxious, in which case we are inclined to say that the appearance of calm is an illusion (*cf.* DSM of Mental Disorders, 5th edition, 189). But must we understand anxiety in this way?

For psychological purposes we will need to take account of visible signs of anxiety, and not merely reports of inner feelings. The visible signs form part of the criteria for suffering from an anxiety disorder, so we also have a notion of anxiety that is, in significant part, defined from an external perspective. In summary the interplay between inner and outer self is often a subtle matter and we should be alert to this kind of complexity.

Important for our appreciation of the self is the concept of a person, which is central to a broad range of philosophical subdisciplines, especially the philosophy of mind, practical reasoning, action theory, political philosophy, and ethics. A central feature of being a person is the capacity, for typical cases, for most of a life, to have thoughts and feelings, make decisions and to have an internal perspective, which will involve seeing oneself as distinct from others and the surrounding world. Persons engage in theoretical and practical reasoning since they not only have a view of what they ought to believe but how they ought to act. These characteristics are important for the role persons play in the political and moral spheres: being held responsible for actions and having rights and responsibilities and so forth. The notion of autonomous action is a central notion in this picture of things (*cf.* Buss 2012; Mele 1992, 2003; Millgram 2015). What has been briefly described here is most charitably described as an idealisation. The notion of an autonomous, free, responsible person that can play a part in a moral community glosses over the complexities and, generally, the messiness of what it is to be a person and how we tell a story about ourselves in a variety of circumstances (Goldie 2012).

In virtue of persons typically having thoughts and feelings and so forth, many philosophers speak of ‘the self’ in a way that appears to be synonymous with ‘person’. This may not be the best terminology, but for the moment let me note that the terminology of ‘self’ emphasises the inner aspects of persons.

3. The Narrative Self

I will look at one way that the notion of a narrative is thought to apply to a self. In this section, I first lay out how we should be motivated to think of the self in narrative form. Then, I show how some philosophers have employed such a notion. Once I have completed these two tasks, this puts us in a better position to understand how self-narratives and the inner and outer self are connected.

What we should understand by a narrative is a matter of some debate, but here are a few suggestions. It is standardly thought to involve a sequence in time. If we focus on narratives such as novels or biographies, then we think of the narrative as something that is evident to a conscious mind and it is so in the mind of the author of the narrative. If we think that a narrative in some way constitutes the self, then, at first approach at least, the self is to be the author of the narrative. An important feature of narratives that is easy to gesture at, but hard to explain in any substantial detail, is the way there is a theme, plot, or narrative arc. It does not contain a mere chronology of events, but there is a connecting thread that gives the narrative a shape or structure. Goldie refers to this kind of structure by employing Paul Ricoeur’s (1984) notion of *emplotment*.

I would like to introduce the bold view that to construct one’s narrative is to construct one’s self, so that the self-narrative constitutes the self. While bold, this view has its contemporary adherents, notably Marya Schechtman (1990, 1996), Alasdair MacIntyre (1984), and Daniel Dennett (1989, 1991). A terminological proposal will permit me to describe this view in a neat fashion. Let us speak of that inner story as ‘the self’, which on this view forms the inner core of the person, where persons are embodied (leaving science fiction cases to one side) as a living human being. That is, the internal narrative that one constructs *is constitutive of* the self. Naturally, the approach has its critics, for example: Peter Lamarque (2004), Galen Strawson (2008), and Simon Beck (2013). In the context of the view in question, this terminological proposal establishes a tight relationship between a self and person. The self is responsible for the inner aspects of a person, or so it would seem at first sight on this view. The human being is a person in virtue of the presence of the self.

One benefit of this way of talking is to allow that while we have a person in virtue of there being a self we can conveniently describe some morally tricky situations. Consider the person with advanced Alzheimer's disease. We may want to say that there is no longer a self, or at least no longer the self that we previously knew. But we still have a person who is owed respect and has rights on the basis of the fact that previously there had been a self that animated that person. Similarly, where we have a newborn child we wish to grant her many of the rights of personhood even though there is no self as yet. It is that potential self, one built into the potentialities of human beings that allows us to treat the new human as a person.

Another benefit of the adoption of the terminology of self and persons is that it enables an easy distinction between the bold view of the role of narratives and this alternative view. Alasdair McIntyre (2003) developed the view that a person lives a narrative. The whole life has a narrative form. For him, no distinction is made between the person and an inner self that has a narrative form. All of the doings and experiences of the person are part of the narrative. In this kind of view a narrative is said to be continuous with life, where we are speaking of the life of a person. In contrast, the view that is my target holds that narratives are discontinuous with life. We can say that because the narrative selects out details from life in the process of constructing that structure, the narrative is thus, of its very nature, incomplete. That is not to find fault with the narrative. On the contrary, it is that capacity to select and emphasise that provides the shape that a narrative offers us. No doubt there are things that could be said in support of McIntyre's notion, but that is not our focus here. We wish to discuss the view that a narrative constitutes the self and it is the self that provides the unifying core of a person.

Now that we have explored the attraction of a narrative to better appreciate what the self is, more should be said about what that notion of narrativity is that may help define the self. The notion of a narrative has been employed in many contexts and there are many attempts at defining it. For example, Noël Carroll (2001, 2007) holds that a causal connection between events is required for there to be a narrative connection between them, even though the earlier event may not be causally *sufficient* for the later one. In contrast, David Velleman (2003) argues that causality is not required for narrativity and instead proposes that it is the emotional cadence in the audience associated with the series of events that is crucial for the structure of a narrative, as this is what makes those events intelligible. He says, "The cadence that makes for a story is that of the arousal and resolution of affect, a pattern that is biologically programmed" (Velleman 2009, 13). Both authors acknowledge that a whole text may not have a narrative structure while containing significant narrative elements. No further purely definitional issues concerning the notion of narrative in general will be pursued here; indeed, a working assumption of this paper is that there may be no single definition that does justice to all common uses (*cf.* Lumsden and Ulatowski 2017). It is plausible that both causal relationships and emotional cadence contribute significantly to the structure of a *self*-narrative and this paper shall proceed to focus on the particular case of self-narratives.

One advantage of a narrative approach is that it suggests a form of holism that unifies a series of experiences. As should be becoming clear, the holism relied on will not be at the level of the whole person. A memory of a certain experience or action is not, unless in rather exceptional circumstances, a discrete item that can be specified completely in isolation. A useful image that can clarify the appeal of holism can be found in David Carr's (1986, 21–30) discussion of Husserl's example of a melody. Hearing someone whistling a tune cannot be reduced to a series of distinct experiences of him whistling particular notes, for each note has the significance it has only in relation to the surrounding notes. In a parallel way, one memory is inextricably bound up with knowledge about other events, people, and places, and indeed the emotional value of the memory depends on just such linkages. What lies behind the appeal to the notion of a narrative in talking about the self is some sense that a narrative provides a structure into which experiences and actions fit, as the analogy with a melody suggests. But it is no trivial matter to specify the nature of that structure, as the dispute between Carroll and Velleman makes clear.

One concern with identifying the self with the self-narrative is the danger that the self-narrative could be totally fictional. Couldn't one's self conception be totally at odds with how one really is? This kind of worry lies behind Marya Schechtman's (1996: 119ff) introduction of a 'reality constraint' into her narrative self-constitution view. In practice, your self-narrative is *not* constructed as a complete work of fiction but has a form that borrows from themes prevalent in the society and others' narratives of you and is based on real world interactions. If the self-narrative is what genuinely is employed to interpret events and guide action, then maybe it is correct to identify it with the self even if it contains serious delusions (Lumsden forthcoming).

The notion of a narrative can be deployed in discussing both inner and outer aspects of the self. The strong view that one's own internal self-narrative defines or even creates the self favours the view that the inner self is the true self. But a narrative approach can be deployed on the other side also. A person's biography, without thinking of an autobiography, will have a narrative form. A biography could scrupulously avoid all mention of the person's inner feelings, at the cost of some sales. I am, of course, thinking here of Douglas Quine's (2020) dryly written biography of his father, Willard Van Orman Quine's life. The biography goes so far as to include places Quine travelled over without landing, or could see from his airplane window. Even in his autobiography, *The Time of My Life* (1985), W.V.O. Quine takes a very sterile view of his life. Again, we should attend to the complexities of the division between inner and outer self. Our self-narratives are, in part, assembled and organised by interaction with others' narrations about us. Any discussion of narrative and the self needs to take on board this feature. The conception of a self-narrative as purely inner is flawed, for we are not purely solipsistic beings.

As already explained, my interest is the bold view that a self-narrative creates the self without being constitutive of it, at least not in a clear cut way. There have been various metaphors for mind, often reflecting the science and the prevailing concerns of the age. For example in the 1990s, the computational mind metaphor was ubiquitous. A narrative notion of the self can be seen in something of the same light, though it is not a notion that comes from science. With any metaphor, for it to be truly useful, it must be taken fairly seriously. For it to earn its intellectual keep, it must not just illustrate some obvious features of the target notion but must be the source of further insights; it should prompt the discovery of previously overlooked features of the target notion that reflect features of the metaphor. On the other hand, we need not, and should not, assume that the metaphor is going to provide insight into everything that is going on. We need to be able to judge when we need to discard the metaphor. Any metaphor is only going to be able to take us so far.

In looking at the narrative self, we should on the one hand identify the ways in which this does provide insights into the nature of the self, but perhaps it is unreasonable to expect that the narrative metaphor takes us the whole way in explaining the self. The subdomains of philosophy whose area focuses upon persons, selves, experience, and action forms such a complex and treacherous intellectual landscape that we should be grateful for any insights that a metaphor such as the narrative one can provide but not be disappointed if it is not going to take us the whole way in conceptualising that terrain. Even so, here is a line of thought that may give the view some initial plausibility, by contrasting a narrative approach to personhood with the more traditional psychological continuity approach.

The psychological continuity view is mainly developed in the specific context of personal identity questions, that is, what it is that determines a person at one time is the same person as that at another time (*cf*. Noonan 2003). A key part of the psychological continuity is the person's memory. That I now can remember acting and experiencing in certain ways in the past is important in determining that I am the same person as the actor and experiencer. The details of the memory theory have been well discussed in philosophical literature since Locke, to whom that approach is attributed, and we should not detain the reader here with the details. Other psychological connections can also be thought important. For example, that I have the same character as a person from before can be relevant to considering us to be one and the same person. Another theme that comes into these discussions is the way that the connections are causal: The second experience is deemed a memory of the first one in part because the first is considered to have caused the second.

A natural way of understanding the nature of the psychological connections is to regard experiences as discrete independent items which can be related by memory and other psychological connections in the way just described. Looked at another way, we could say that this approach to personal identity is reductionist. That is, what it is to be a person reduces to the mental states that stand in the prescribed relationships. But can we rely on such discrete mental states to form the basis of our theory? In many cases the experiences recorded in memory are inextricably connected with other experiences. Suppose I recall going to a dinner party with some friends. That bare memory, or a richer one which incorporates a series of events and conversations that occurred at the dinner party, will not make sense without some kind of backstory about the friends, roughly where they were living at the time and various other supporting details. Thus there is a form of holism about memory. Discrete memories are not plausibly the building blocks of a person. That is not to say there could not be a discrete memory. I might remember attending a dinner party where a guest literally fell asleep in his soup, while that is not supplied with any significant surrounding context. There would need to be some background understanding of dinner parties, soup, and falling asleep, but, in the situation I am supposing, there are no specific connections with individuals or a location. Consider how rare and strange such a memory is. We don't plausibly build a person from such materials as that.

Here is where the notion of a narrative might assist us. A defining feature of narratives is that they have a form or structure into which particular incidents fit. This is one way of capturing a form of holism amongst incidents, which in the context of providing an account of selfhood or personhood expresses a holism amongst experiences in memory. Thus a memory is not free-standing but is the memory it is only in the context of the narrative in which it appears. When Goldie speaks of the process of *emplotment*, he speaks of there being four parts, one of which is the raw material. This might be taken to suggest that Goldie is leaning towards a reductionist approach, where the raw material is composed of discrete experiences. Now there may be an element of truth in that, but, significantly, Goldie insists that the descriptions of persons, actions, events, and objects that comprise the raw material are 'rich', that is they include 'thoughts, feelings, moods and emotions, character traits and personality'. One significant aspect of this richness is the inclusion of the emotional dimension, which is something not mentioned so far. Yes, surely that is important both for narratives in general and for memories that are part of a person's self. But that is not the only significant aspect of this richness. Note the raw material includes descriptions of persons. Thus in my example of recalling having dinner with friends, the memory of the event is inextricably linked with knowing something about the guests (though of course one might have forgotten exactly who was there). Then we include character traits and personality. Again, this takes us beyond the bare snapshot of the event. That John and Mary argued about a certain topic made sense and thus was memorable on the basis of one's understanding of their respective characters. For these reasons, I see Goldie as explaining, in somewhat different terms, how there is a holism in memories that are significant for being a person.

4. Plurality of Narratives, Plurality of Selves

Here's a subtle truism about the self: *a self typically lacks the kind of coherence that a single narrative would be expected to provide*. This disunity or plurality of the self is exacerbated by an examination of someone's outer self. An extreme illustration of such disunity may be provided by *Six Days of the Condor*'s Ronald Malcolm, a CIA agent whose clandestine services include nothing more than 'reading books'. Codenamed "Condor", Malcolm returns one fateful day from lunch to find everyone in his office assassinated. Subsequently, he contacts headquarters to report that his office has been hit but discovers that he cannot trust anyone in the organisation. One minute, Malcolm is enjoying a hoagie, and the next, he is fighting to remain alive. At numerous points, Malcolm himself has difficulty reconciling the life he led before the 'hit' with the life he leads following it. This fictional example is an extreme one, so the focus should be on the variety of narratives that can be involved in a much more mundane life.

The self, expressed in narrative terms, is not unified in a strong sense but disunified such that there may be no single narrative that unifies the whole life of a person, as, for many of us, there are various relatively discrete roles that we play in different settings, which are captured by different narratives. There need be no monolithic and overarching narrative that draws together the person's myriad component narratives. When we are at university lecturing to a crowd of students, our disposition and how we interact with them is distinct from how we behave at an intimate dinner with our partner. The lecturing situation is part of a narrative thread that includes one's history of employment and professional obligations; the coherence of this thread, in turn, relies on details related to expectations of how an instructor ought to behave in that situation, expectations of the role of students, and a history of teaching experience, all of which provide guidance as to how to read situations and react to them. In contrast, private moments with a partner are embedded in a narrative that does not relate to professional obligations but to a personal common history and emotional ties or conflicts. We can say that within the person there are different narratives running in the two situations, where each story provides not only a distinct history of events but also a distinct set of themes, vocabularies, descriptions, values, and associated emotional colour. The different patterns of behaviour follow from those different stories. Call each story a 'narrative thread' (*cf.* Lumsden 2013; Lumsden and Ulatowski 2017).

The contrast between two narrative threads just discussed rested on two different environments, but it would be a mistake to think that the environment completely determines the narrative. The subjective position and interpretive lens matter for how the narrative unfolds. This is evident when we consider how the very same event could be interpreted differently by distinct people who approach the situation on the basis of different narratives. Compare this with how different witnesses provide very different accounts of events such as traffic accidents or barroom brawls. To provide a somewhat stylised example, consider the lawyer, the mechanic, and the novelist who witness a traffic accident. While there is likely to be some overlap in the descriptions provided by the three, each might attend to different aspects of the accident: The lawyer attends to the issue of fault, the mechanic attends to the vehicles' damage, and the novelist attends to how the accident fits into the lives of the people affected. For each of them, that one incident fits into a pattern of events that is captured in a narrative that they, but not the others, possess. This shows how different narratives can be brought to the same situation, though in that case, those different narratives are themselves grounded in the different situations that comprise the working lives of the three different observers.

To take a narrative approach to the self locates unity and disunity within a rich set of materials. A narrative of the kind that concerns us here involves a history of events incorporating a stream of experiences replete with an emotional charge and a set of values, and which tend to be indicative of certain character traits. Thus, each narrative thread will have its own unique flavour or perspective. The irritated commuter, let us suppose, has experienced a long series of adverse commuting incidents over months or years: delays, misinformation, overcrowding, and inadequate heating or cooling. The emotional reaction and pattern of response to these, whether of frustration, resentment, or anger, provides the source of the commuter's action in the current circumstances, such as accusing an employee of incompetence.

While the example of the irritated commuter highlights emotional reactions, rational thought and plans and policies are very likely part of the narrative too. The commuter's memory of previous attempts to respond to a train cancellation can lead to a new strategy to reach the destination. The use of narratives can help explain how actions flow from a rich structure that includes, but is not limited to, the formal character of practical reasoning. This is in line with Peter Goldie's (2009) cautious discussion of the way that our employment of narrative in planning future actions incorporates character, personality traits, and emotional dispositions. He opposes Bratman's (2007) view that rests fundamentally on higher-order self-governing policies.

My view is that such a unity of inner self holds only *within* each narrative thread. Even so, we fully acknowledge connections and coordination among different narrative threads, which can ensure they do not interfere with each other and do allow the person to function in the world. Might plans and policies provide that sort of connection? Indeed, plans and policies may provide some connections between narrative threads, but we should not assume there is a singular form of connection.

5. Taking the narrative metaphor too far

The narrative metaphor captures a form of holism amongst different aspects of a person's life, which aggregates to what the self is. While there are component memories and experiences, those things are not discrete items that can be bundled together to form a self. Rather, a narrative explains how each experience, and a memory of an experience, is what it is only in relation to other aspects of one's mental life. So to take the metaphor to its logical conclusion we might reasonably think that the entirety of a self is a complete narrative. The idea sounds simple.

There is reason to doubt the plausibility of this metaphorical model. It misses the multifaceted character of a life. Suppose we grant that experiences are the composition of a narrative self – need we assume that the different parts of a person's life be bound together into a neat and tidy narrative form, much as one might come to expect of a novel or biography? There could be a form of narrative cohesiveness amongst the rather disparate events in my work life that may not connect with events in my home life or the events related to my primary hobby, like Malay Zapin Api line dancing. Even that could be too simplistic. Within my work life, for example, there may be different narrative threads, where the personnel are different and the events and emotions are quite disjoint. Add to that my social life, which includes a group of friends from primary school. There could be a number of narrative threads in a person's life that do not connect in any significant way. There may be no 'meta-narrative' thread that binds all these narrative threads together. Few of us have two quite distinct personalities like Dr. Jekyll and Mr. Hyde, but many of us move in different settings that make very different demands on us and prompt different styles of behaviour. It does not follow that the points of contact need be so rich as to provide a meta-narrative thread which sustains a narrative unity.

If this observation is correct, then in defining the self we could echo Hume and say that the self is a bundle of narrative threads. To take this stance is to find a certain role for narrative but not to assume it provides the whole answer. To speak of a 'bundle' here could suggest that the group is less than a perfect unity. That indeed may be the best approach to the self. There is a kind of functional unity, because usually one of the threads is centre stage at any one time. We could play along with this image and suggest that other threads may hover at the edge of the stage or be waiting in the wings. The unity of a normally functioning person in normal circumstances consists in the fact that there is a relatively smooth handover from one thread to another. In fact, on occasion a pair may need to operate in tandem, switching control back and forth as the need arises. In this ability for narrative threads to swap control appropriately is the unity of the self, such as it is.

Perhaps this picture sounds a little extreme. It does not seem counterintuitive to suppose that surveying aspects of my life *just is* a matter of contemplating the whole life narrative. Such a thing may be possible, but note how that activity is an unusual and very peculiar activity. The cognitive load it places upon the person to survey not only what aspects of life lay before her now but also how these aspects fit properly into a wider narrative whole is quite large. The normal level of functioning when it comes to identifying and disseminating what the self is is more of the nature of control switching. This view of things can take comfort from Elijah Millgram's work on action, especially his (2015). There, he emphasizes that it is a critical fact about humans that we are serial hyperspecialisers. It is part of the way we are adapted to the world that we can focus on particular things for reasonably extended periods, and then can move on to something else.

6. Conclusion

Recall the earlier discussion concerning the complex interplay between the inner and outer self. A narrative view might seem to suggest that the real self is a purely inner narrative. This would be a mistake. What has instead been argued here is that a narrative which is constitutive of a self balances the inner self—composed of one's perceptions, emotions, and impressions—with the outer self—composed of differing narrative threads that find their origin in the roles that a person plays in work and life. Such a view recognises and appreciates the interplay between different aspects of one's life. One minute we are parents to our children and the next we are on a Zoom call with our colleagues in Nairobi. The narrative autobiography that encapsulates the inner and outer self may be a plurality of disunified selves, but it most accurately captures what is constitutive of the self.

Acknowledgements

I am grateful for conversations with David Lumsden, Dan Weijers, and Doug Campbell that led to a draft of this paper, and I am especially thankful for the generous financial support of the University of Connecticut Humanities Institute that provided me with valuable time and space during a visiting residential fellowship to help complete the project.

References

Beck S (2013) The Misunderstandings of the Self-Understanding View. *Philosophy, Psychiatry and Psychology* 20(1): 33–42.

Bratman M (2007) *Structures of Agency*. Oxford: Oxford University Press.

Buss S (2012) Autonomous Action: Self-Determination in the Passive Mode. *Ethics* 122(4): 647–691.

Carr D (1986) *Time, Narrative and History*. Bloomington, IN: Indiana University Press.

Carroll N (2001) On the narrative connection. In: *Beyond Aesthetics*, New York: Cambridge University Press.

Carroll N (2007) Narrative closure. *Philosophical Studies* 135: 1–15.

Dennett D C (1989) The Origins of Selves. *Cogito* 2: 163–173.

Dennett D C (1991) The Reality of Selves. In *Consciousness Explained*, Boston: Little Brown and Company: Chapter 13.

Descartes R (1637/1998) *Discourse on the Method of Rightly Conducting One's Reason and Seeking the Truth in the Sciences.* [In:] *Descartes: Selected Philosophical Writings, translated by John Cottingham, Robert Stoothoff, and Dugald Murdoch, Cambridge: Cambridge University Press,* 20–56.

Descartes R (1641/1998) *Meditations on First Philosophy*, [In:] *Descartes: Selected Philosophical Writings*, translated by John Cottingham, Robert Stoothoff, and Dugald Murdoch, Cambridge: Cambridge University Press, 73–159.

Descartes R (1644/1998) *Principles of Philosophy*. [In:] *Descartes: Selected Philosophical Writings*, translated by John Cottingham, Robert Stoothoff, and Dugald Murdoch, Cambridge: Cambridge University Press, 160–212.

Diangnostic and Statistical Manual of Mental Disorders, 5th edition. (2013) Washington, D.C.: American Psychiatric Association.

Flanagan O (1991) *Varieties of Moral Personality: Ethics and Psychological Realism*. Cambridge, MA: Harvard University Press.

Flanagan O (1996) *Self Expressions*. Oxford: Oxford University Press.

Frankfurt H (1988) *The Importance of What We Care About*. Cambridge, UK: Cambridge University Press.

Goldie P (2009) Narrative thinking, emotion, and planning. *Journal of Aesthetics and Art Criticism* 67(1): 97–106.

Hume D (1739–1740/1973) *A Treatise of Human Nature*, ed. L.A. Selby-Bigge. Oxford: Oxford University Press.

Kripke S (1982) *Wittgenstein on Rules and Private Language*. Oxford: Blackwell.

Lamarque P (2004) On Not Expecting Too Much From Narrative. *Mind and Language* 19: 393–408.

Locke J (1700/1975) *An Essay Concerning Human Understanding*, 4th Edition, ed. Peter H. Nidditch. Oxford: Clarendon Press.

Lumsden D (2013) Whole Life Narratives and the Self. *Philosophy, Psychiatry, and Psychology* 20: 1–10.

Lumsden D (July 2015) Narrative construction of the self and the accuracy problem. In: *Mind, Body and Self*, Proceedings of the Society for Philosophy and Culture Conference, Wellington.

Lumsden D, Ulatowski J (2017) One Self Per Customer? From Disunified Agency to Disunified Self. *Southern Journal of Philosophy* 55(3): 314–335.

MacIntyre A (1984) *After Virtue: A Study in Moral Theory*, 2nd edition. Notre Dame, IN: University of Notre Dame Press.

Mele A (1992). *Springs of Action: Understanding Intentional Behavior*. Oxford: Oxford University Press.

Mele A (2003) *Motivation and Agency*. Oxford: Oxford University Press.

Millgram E (2015) *The Great Endarkenment*. Oxford: Oxford University Press.

Noonan H (2003) *Personal Identity*, 2nd edition. London: Routledge.

Parfit D (1984) *Reasons and Persons*. Oxford: Oxford University Press.

Quine D B (2020) Willard Van Orman Quine, 1908-2000: Philosopher and Mathematician. https://www.wvoquine.org.

Quine W.V.O (1985) *The Time of My Life*. Cambridge, Mass: MIT Press.

Schechtman M (1990) Personhood and Personal Identity. *Journal of Philosophy* 87: 71–92.

Schechtman M (1996) *The Constitution of Selves.* Ithaca, NY: Cornell University Press.

Strawson G (2008) Against Narrativity. In *Real Materialism and Other Essays,* 180-227. Oxford: Clarendon Press.

Velleman J D (2003) Narrative explanation. *Philosophical Review*, 112(1): 1–25.

Velleman J D (2009) *How We Get Along*. Cambridge, UK: Cambridge University Press.

CONCEPTUALIZING THE SELF: THE ROLE OF NARRATIVES

Hari Narayanan V
Department of Humanities and Social Sciences
IIT Jodhpur, India
e-mail: hari@iitj.ac.in

Abstract:
The paper seeks to argue that different ways in which the self is understood, even if radically distinct from one another, are cases of different narratives. This is done by appealing to conceptual metaphor theory. The paper begins by briefly explaining the difference between the minimal and narrative self and then argues that even radically different ways of understanding the self are cases of different narratives arising out of a metaphorical understanding of abstract concepts.

Key words: narratives, self-conception, metaphors, minimal self, narrative self.

Introduction

The importance of narratives in human life is aptly captured in the oft-quoted statement of Daniel Dennett, “Humans extrude narratives just like beavers build dams or spiders spin webs” (Dennett 1989). That is to say, it is quite common and instinctual for humans to engage in inner speech that gives rise to the rich story that we carry about ourselves. This capacity is clearly linked to the common tendency of thoughts to wander from one content to another, which is famously called the stream of consciousness. The stream of consciousness is, in the words of William James, “like a bird's life, made of an alternation of flights and perchings” (James 1890). That means mental contents, like the flight of a bird, continuously change. But even if mental contents shift from the memory of something that happened a decade ago to a projection of what is likely to occur next year, most spontaneous thoughts are centered around the self in one way or another.

A distinction is often made between the minimal and narrative self (Gallagher 2000; Gallagher & Zahavi 2012). The former is intrinsic to any conscious state by virtue of the fact that we are the subjects who are conscious of it. That is, when we experience anything, whether it be an external perception of an object or an internal perception of something like pain, the experience is always *for* a subject. We do not have to make any inference to know that we are the ones who undergo that experience. In other words, experience always presupposes an experiencer. This is minimal because it does not leave room for any detailed account of what the self is, of its traits, or of its life story. Gallagher defines minimal self as follows: “Even if all of the unessential features of self are stripped away, we still have an intuition that there is a basic, immediate, or primitive 'something' that we are willing to call ‘self’” (Gallagher 2000). This claim is similar to what is called the ubiquity thesis, according to which some form of self-consciousness is present in all kinds of conscious experience (Milliere & Metzinger 2020). This kind of minimal self can be the direct result of bodily processes like interoception and proprioception. As Anil Seth observes, representations of physiological states are associated with the basic, pre-reflective form of selfhood (Seth 2013). In fact, there have been attempts to show how interoceptive processes underlie our complex mental life (Tsakiris & Preester 2019). It could well be the case that such a primitive form of self-consciousness is present throughout the animal kingdom (Bermudez 2018).

Apart from maintaining the balance of internal states, an organism has to constantly consider its boundaries while successfully navigating its environment, a conscious process which can underlie the minimal self. In this sense, the minimal self is both embodied and always present in the world (Gallagher & Zahavi 2012). This minimal self is non-conceptual because having a minimal self does not presuppose the ability to entertain concepts as ordinarily understood. It can be a function of non-conceptual first-person content because ascription of that content to a being does not presuppose the ability of that being to engage with phenomenal content at the level of abstract concepts (Bermudez 1998). In this sense, the minimal self is pre-reflective because it does not presuppose any ability to engage in reflection through concepts.

But in folk psychological understanding, the self is primarily an entity that has an identity, traits or characteristics, and a life story to tell. The notion of the narrative self captures this aspect of selfhood. If the minimal self is invariably present in any conscious experience and can be more or less similar across different human beings or even other species, it can be distinguished from the narrative self whose contents vary from person to person or even within a single person across different periods of time. The narrative self arises with the human ability to form a complex conceptual network. If the minimal self appears as essential to activities necessary for basic survival, then the narrative self is the result of higher-order mental processes, namely, the ability of the language-using organism to form abstract concepts about many things, including oneself. That is to say, the minimal self is more a result of direct bodily experience, but the narrative self stems from the ability to abstract and maintain a distance from the immediate environment.

The narrative self consists in the ongoing life story of a person. It is rich in meaning and comprised of the experiences that one has had as well as one's values, aspirations, and other features of identity that collectively constitute what is commonly called the person. It involves both endorsement and rejection of many things based on a consistent worldview or set of values, without which it would be impossible to lead a life in conformity with a cultural and/or personal framework that brings order, understanding, and meaning to one’s life experiences. Hence, it is the narrative self that underlies aspects of human life that go over and above biological survival. Here, it is important to note that the word ‘narrative’ must be understood in a broad sense. Whenever an assertion about the nature of the self takes the form of a verbal pronouncement, such as, “I am this,” the narrative self is in operation. Though the process of offering truth-apt statements about the self presupposes linguistic ability, this does not mean that ‘narrating oneself is always achieved through explicit linguistic utterance. Rather, narratives may be unspoken, unconscious, or implicit. However, in all of these instances, the content of the narrative self is conceptual (as opposed to merely perceptual) in nature because selfhood can be understood through this higher-order narrative form only when an organism is capable of constructing a coherent account of experience through concepts.

To possess concepts is to be able to engage in making inferences because conceptual systems can be understood in terms of a network with strong interconnections. Narratives, as commonly understood, involve such inferences, and therefore it is reasonable to hold that narratives are primarily a matter of conceptual activity even if they are not always reported through linguistic utterances. It is pertinent to note that this line of thinking is within the framework of the approach adopted in cognitive linguistics according to which conceptual structure is prior to and richer than language (Green & Evans 2006). Language, as per this account, prompts or triggers the exercise of this capacity for conceptualization, resulting in understanding and meaning-making.

Thus, while a newborn infant lacks any narrative self, it is constructed over the course of ontogenetic development. The narrative self, as per this account, presupposes the minimal self. That means the ability to have concepts about oneself and form complex narratives requires non-reflective understanding of oneself. As Bermudez argues, full-fledged linguistic self-consciousness is grounded in non-linguistic self-consciousness (Bermudez 1998). And as Gallagher and Zahavi put it, experience has to be given as 'mine' before it can be united with other content to form a life story (Gallagher & Zahavi 2012). In this way, the minimal self is more like the background on which different portraits are painted through narratives. When portraits are salient, the background is hardly noticed. By analogy, the minimal self can never be present in consciousness as an entity that can be described, given its non-conceptual content, while at the same time, it remains an essential (though unarticulated) presupposition before any conceptual claims about the self can be offered.

The above account gives rise to two major questions. Given the universality of the minimal self, why is it the case that the narrative self varies widely from culture to culture and even within individuals over time? Answering this question requires considering many outstanding issues, such as the exact nature of the relation between the minimal and narrative self (A. Bortolan 2020), and the precise role of narratives in selfhood (Dennett 1992; Taylor 1989; Schechtman 1996; Strawson 2004). It is beyond the scope of this paper to offer a detailed discussion of these issues. I shall restrict myself to answering the following question: Though the minimal self is universal, how is it the case that the narrative self varies radically across cultures? Appealing to conceptual metaphor theory, I argue that substantively different ways of understanding the self are cases of different narratives. In the next section, I shall briefly discuss the different ways of conceptualizing the self. In the subsequent part of the paper, I shall introduce conceptual metaphor theory and argue that narratives formed out of different metaphors can give rise to radically different ways of understanding the self.

Varieties of Self-Understanding

How does the self manifest implicitly and automatically in everyday behavior, and thus, how should it be conceptualized? Since the minimal self is pre-conceptual, this question applies to the narrative self only. (Hereafter, the term 'self' is used to refer to the narrative self.) Now, to be precise: The question I am asking is concerned not with determining what the metaphysical nature of the self actually is but with sussing out how it appears in the way we conduct ourselves. Even if there is no monolithic way of conceptualizing the self, it is reasonable to hold that in the modern world, which increasingly subscribes to an ideology of individualism, the self is understood as an entity that is distinct from the body and that sustains and enriches itself in various ways. Miri Albahari provides an interesting account of the way we automatically understand ourselves. According to her, four factors underlie human self-conception: autonomy, consistent self-concern, ownership, and this-ness (Albahari 2006). Autonomy refers to the feature of free will according to which human beings have the freedom to choose between various options before making a decision; this makes them responsible for their decision. That means the self is understood to be something that is endowed with free will. The feature of consistent self-concern refers to the tendency of thoughts to be centered around the 'me'. Ownership refers to the tendency to consider objects and even other beings one's possessions. This-ness refers to a belief in the uniqueness of the self, a uniqueness which is over and above any special physical and mental traits that characterize the individual. These four features, according to Albahari, give rise to a thick psychological boundary of the self with the world.

This thick psychological boundary underlies what is called the free-floating or separative self. The boundary is psychological because it bears no connection to the physical limits of the biological organism in space. It may be recalled that the minimal self is closely associated with the body's internal and external aspects, while the narrative self is the result of conceptual capacity. As far as the body is concerned, there is indeed a boundary with the rest of the world, but this is different from the psychological boundary discussed here. To better understand the nature of this psychological boundary, consider one of its previously enumerated features: free will. Belief in free will is belief in the subject's ability to act in a manner that appears to be above the causal order of the physical world insofar as the effects produced in part through human acts would not have been the effects produced by a purely mechanistic system in which no actors participated. When we say that we have acted out of free will, we are rejecting the possibility that our actions are simply part of the complex causal network present in nature.

Similarly, consider ownership. A person can feel a sense of ownership over many things, including land, money, objects, and even other persons. There are institutional arrangements to ensure that individuals not only perceive themselves as owners but are invested with the rights and responsibilities of owners in cases in which ownership is considered lawful. The tendency to own things or to feel possessive of them is quite common and can be understood as a process of identification; that is, the self wants to identify itself with many other things and expresses this through the relation of ownership. The relation of ownership crucially points at the thick psychological boundary the self has with the rest of the world since the dyadic relation between the self and the object with which it identifies can only exist if boundaries demarcate the self and the thing owned. Recall that the self is called free-floating, and this is because its identifications keep changing. At one point ‘the I’ may be identified with a man who was praised by his colleague the other day, but in the next moment ‘the I’ becomes identified with someone who is anxious about a meeting with the boss the next day. Similarly, the things that we want to possess (and thus to identify ourselves with) continuously change. The shifting nature of self-identification points to another quality that makes the self free-floating: its separative nature. It considers itself to be separated from the rest of the world with no feeling of abiding connection to it, except perhaps to the things it considers as its own. Human alienation or estrangement from nature is common in the Industrial world, as is the tendency to look at nature as something to be controlled or conquered. The sense of being alienated and separate from the rest of the world with only limited integration through the relation of ownership follows from the thick psychological boundary of the self with the world.

Have humans always conceived of themselves in the way characterized above, and has the self always maintained a thick psychological boundary with the world? Have humans always had a tendency to feel alienated or estranged from nature? There may be no way to provide a definite answer to these questions, but there are indications to the effect that the way humans understand themselves has varied significantly throughout our history, and the particular model of self-conception discussed above is currently the most dominant one but is not the only one.

As evidence of this, consider the distinction often made between individualism and collectivism (Hampton & Varnum 2017). According to the individualist model, the self is independent and distinct. For cultures in which individualism represents the dominant ideology, traits like goal-orientation, ambition, and assertiveness are highly valued. By contrast, a collectivist way of looking at the self values interdependence more than independence. As such, belonging to a community or close circle of people is a more important aspect of self-conception than one’s distinct individuality, and traits like affection and cooperation are more highly valued than ambition and assertiveness.

The individualist and collectivist ways of looking at the self may not appear radically different because both types are present in our time. But it is clear that they are distinguished by the degree to which the four factors of the modern individualistic self are present. For instance, the feature of autonomy, and thus the belief in freedom of the will, is likely to be less salient in collectivism, which emphasizes values like interdependence and cooperation. It has been claimed that the modern concept of the free will began with St. Augustine, and that prior to the introduction of this concept,notions of freedom and responsibility were separated from that of the will (Sorabji 2004). The notion of free will continued to gain in prominence, becoming central to the project of self-understanding during the Enlightenment period. Enlightenment thinkers popularized the idea that all human beings are equal because they all are free and can therefore use reason to guide their conduct and change the world around them. A similar argument can be made about ownership as well. Private property has become common and acceptable only in the last couple of centuries, reflecting a trend toward greater technological and material innovation and progress (Morgan 1877). Prior to this cultural shift, people tended to express greater interest in spiritual matters than material ones; what awaited one in the afterlife was of more importance than the material property one owned in the here and now, and hence, the relation of ownership was far less likely to significantly impact one’sself-conception. Given this historical fact, it appears true that both autonomy and ownership are less salient in the collectivist way of looking at the self.

In line with this perspective, some theorists claim that human beings used to have a substantial-ly different way of conceptualizing themselves in ancient times. One such claim is Julian Jaynes' bicameral hypothesis. Though there is no clinching evidence to prove the hypothesis, Jaynes' ideas are worth considering. The crux of Jaynes' claim is that until around 3,000 years ago, all human beings had a radically different way of understanding themselves (Jaynes 1976). According to Jaynes' hypothesis, this is because the human mind was divided into two parts, the commanding and the obeying part, which accounts for the bicamerality. When humans encountered situations in which habitual reaction was not possible, they 'heard' voices and acted accordingly. They believed the voice came from outside when in fact they were perceiving what we now refer to as the inner voice. According to Jaynes, these humans lacked a conception of inner space, and therefore, they could not figure out that the voice was not from an external source.

Some anthropologists have claimed that in some cultures there exists a different kind of mentality. Jean Smith (1981) claims that the self in Western culture is conceived of as the driver of a vehicle, whereas in Maori culture, it is akin to a passenger only, observing experience rather than controlling experience. It is common for the Maori to use expressions like the "mouth speaks" or the "hand takes", which presumably indicates an absence of belief in the agential self. The feature of autonomy is largely absent in such a self-conception, and the individual is not seen as an agent who shapes their own destiny. In Smith's words, "it was not self which encompassed experience, but experience which encompassed the self."

The case of the Dinkas of Sudan appears more radical. Godfrey Lienhardt (1961) points out that Dinkas do not make any distinction between the inner world and the outer world. According to him, they do not have a concept of mind similar to the one we employ. The concept of mind is normally employed to explain the inner origin of behaviour or serve as the theoretical locus of experience. But Dinkas do not use categories like memory and belief to account for their behaviour. For instance, if a person saw something and his behaviour changed afterwards, the Dinkas would claim that the thing continued to follow him, resulting in the altered behaviour. This explanation does not include the concept of memory as an internal process. Similarly, since a person's behaviour is always understood to be shaped by external forces, this does not leave any room for a robust account of the inner realm as a source of behaviour. This, in turn, suggests a conceptual scheme that does not embrace the concept of mind itself since conceptual schemes that do include the concept of mind typically assign to it a primary role in explaining behaviour. In consequence of this, the world, for Dinkas, is not an *object* of study because there is no subject (self) that stands sharply separated from the world. Such a view is antithetical to the modern self-conception constituted by a thick psychological boundary with the world.

It used to be common to ascribe many aspects of inner life, such as dreams and creative acts, to external sources, and only in recent times has the conception of the self as fully responsible for all inner processes gained wide acceptance (Fox & Christoff 2018). Across cultures, dreaming, for example, has been attributed to external sources such as divinity (Kracke 1991 as cited by Fox & Christoff 2018). Some cultures used to consider dreams a direct window into supernatural reality. Other cultures used to hold the view that dreams were the result of the wandering of the soul during sleep. Similar claims have been made about artistic creations and other achievements. As Darrin McMahon points out, the notion of genius was long construed in religious terms (McMahon 2013). If the concept of a full-fledged inner world in which thoughts, feelings, and dreams appear from within is relatively recent, then it follows that the phenomenon itself may be recent. In this case, there is nothing implausible about the bicameral hypothesis and similar claims about a radically different kind of subjective life.

What the above discussion demonstrates is that the ways in which the self is conceptualized form a spectrum ranging from an independent, autonomous self to one in which it is not even in control of experience. How is it possible that such different ways of understanding the self arise? In the next part of the paper, I shall try to explain this variation in terms of conceptual metaphor theory, which, in turn, demonstrates the role of narratives in the formation of self-conception.

Conceptual Metaphor Theory

The fundamental tenet of conceptual metaphor theory is that abstract concepts are understood metaphorically in terms of concepts that are directly grounded in bodily experience. The concept that is understood metaphorically is called the 'target' and the concept used to understand the target is called the 'source' (Lakoff & Johnson 1980, 1999). The theory has its basis in the observation that examples of understanding one concept in terms of another abound in everyday language. For instance, when we say, "his smile is warm", we understand the concept of affection in terms of warmth. Similarly, when we say, "prices rose", we understand the concept of increase in quantity in terms of vertical expansion. Or, when we say, "he is a close friend", we understand intimacy in terms of physical proximity.

In all the above examples, it is clear that the concept used to understand another concept is concrete and can be derived from direct bodily experience. Warmth is a matter of bodily sensation, and vertical expansion and physical proximity are qualities of external reality that are directly perceived by embodied subjects. Further, there are reasons why a particular concept, say, intimacy, is understood in terms of physical proximity. In childhood, intimate people are normally physically proximate and this establishes a connection between these two concepts. Similarly, affection is felt in early childhood when being held closely, which, in turn, results in bodily warmth. This line of thinking leads to an embodied understanding of meaning and concepts because even abstract concepts can be traced to the concrete bodily experience that all humans normally participate in by virtue of their biological make-up (Johnson 2007).

The notion of an image schema is central in explaining the process of concept formation through bodily experience. Image schema, as defined by Gibbs and Colston, are “dynamic analog representations of spatial relations and movement in space” (Gibbs & Colston 1995). Bodily experiences that are common to every child give rise to the image schema of concepts like 'on', 'above', 'toward', etc. The body itself is vertically oriented, and it is also common to see things that are kept on top of other things; these basic perceptions can give rise to the concepts of 'up' and 'down'. The experience of things (e.g. food) going inside the body can give rise to the container schema, which includes the concepts of inside, outside, and a boundary. The experience of moving towards something may also give rise to what is called the source-path-goal schema.

The schema generated at the pre-conceptual level are used in understanding concepts that are abstract. Once the container schema is present it is possible to understand many complex concepts by applying it (Johnson 1987). For example, making sense of the statement, "I fell in love", involves understanding the concept of love in terms of a container (that one can metaphorically fall into). Similarly, when we say, "the economy is in trouble", we understand trouble, like a container, as something that encompasses or encloses. Finally, when we understand life in terms of a journey, it is the source-path-goal schema that is used.

The schema of an object is easily generated in childhood through constant exposure to multiple objects, including one's own body. When we understand an object, we are able to make a distinction between the object itself and its properties. This is possible because of the part-whole configuration of an object. Many of the abstract concepts we use, such as justice and truth, may be understood in terms of concrete objects, allowing us to apprehend their qualities with greater ease. The common symbol of the scales of justice, for instance, gestures at the balancing and adjudicating aspects of justice through a familiar and comprehensible object. Certain basic schema developed by means of bodily experience and contact with the material world underlie the abstract concepts we understand in our everyday life. The way we understand abstract concepts is thus constrained by the kind of bodies we have and the kind of experience we participate in.

It is important to note that there can be multiple metaphors operating that allow us to understand an abstract concept . Metaphors obscure certain aspects of a phenomenon while highlighting certain other aspects, so the metaphor we use may determine the way we understand a concept in important ways (Lakoff & Johnson 1999). For instance, when we understand life in terms of a journey, the aspect of purpose in our actions gets highlighted. If we instead understand life in terms of a drama, purpose becomes hidden while conflict may be emphasized. At the same time, if we avoid metaphor, our understanding of abstract concepts may remain impoverished. Metaphors make our concepts richer by enabling complex inferential relations. So, when time is understood as a ‘resource’, several inferences follow, including the notion that we should spend time wisely and that there is a scarcity of time. The complex conceptual network thus generated underlies the vast and intricate manner in which human life is structured.

There have been attempts to find experimental evidence in support of the embodied approach to the origin of concepts and meaning. Consider simulation semantics, which holds that whenever we encounter a word or sentence its meaning is understood by means of simulating the experience that the linguistic item describes (Bergen 2012). This means that the neural regions responsible for sensory-motor activity and emotions play a role in linguistic understanding as well.

Before going further, it is important to briefly discuss the importance of simulation in human mental activity. As Bergen points out, simulation plays a big role in everyday life. For instance, when we visualize a landscape or imagine hearing sounds that aren't there, simulation is at work. We can also simulate actions by imagining that we are performing them without actually moving at all. These kinds of mental events are very common, and it may well be the case that many of them occur at the unconscious level. Several types of evidence, such as fMRI, electrical stimulation, and psychological studies, support the hypothesis that the same neural regions are involved both in performing a task and in simulating or imagining it (ibid pp. 26-39).

We can easily understand and even visualize concepts like "golden mountain," even if we have never experienced them in real life, precisely because we can simulate what it is to see both golden colour and the form of a mountain. This is because meaning is not merely a matter of some fixed definition in a dictionary but is informed by our rich and varied sensory experiences. Much of this simulation goes on automatically at the unconscious level, so subjects may be unaware that they regularly simulate experience and thus unlikely to report it. Nonetheless, studies show that merely mentioning an object elicits visual stimulation. For instance, in a study conducted by Rolf Zwaan (as cited by ibid pp. 53-54), subjects were given either of the two sentences, "The carpenter hammered the nail into the floor" or "The carpenter hammered the nail into the wall." After the subjects had read either of the two sentences, they were shown two pictures, one in which the nail is pointing downwards and the other in which the nail is pointing sideways. It is commonly known that when a nail is hammered into the floor it will point downwards and when it is hammered into the wall it will point sideways. After being shown the pictures the subjects were asked whether the picture had been depicted by the sentence. Subjects who had read the 'floor sentence' were able to identify the downward pointing nail as related to the sentence they had just read more quickly than members of the other group, and those who had read the 'wall sentence' were able to identify the sideways pointing nail in a similar fashion. This result supports the hypothesis that we understand the meaning of sentences by simulating the scenes that they depict.

The simulation hypothesis extends to metaphorical understanding as well. In a study, subjects were prompted to perform an action, either grasping or swallowing (ibid p. 199). Then, they were exposed to a metaphorical phrase, such as 'grasp a concept' or 'swallow your pride.' It was found that those who were prompted to perform the action of grasping could understand the metaphorical phrase 'grasp a concept' more quickly than those who performed the action of swallowing. Similar results were recorded in the case of subjects who performed the action of swallowing when they were shown the phrase 'swallow your pride.' This indicates that understanding abstract concepts involves bodily simulation.

The narrative self is an abstract concept because it is over and above the body, which is concrete and tangible. If abstract concepts are metaphorically understood, as conceptual metaphor theory contends, then what are the basic concepts that we use in understanding the self? Can we understand the different ways in which the self is conceptualized in terms of different metaphors? If so, we must acknowledge that g since metaphors enable complex inferential relations, they give rise to multiple narratives, and thus the self must be the sort of thing that cannot be identified with a single authoritative narrative.

Narrating the Self

The self is certainly understood in terms of its nature as an object, and many of the differences we notice between various approaches to self-conception can be explained by the different properties attributed to the object called self. Some thinkers contradict this claim, arguing that the self is quintessentially a subject and therefore cannot be understood as an object. t Since the present paper does not tackle the metaphysical question of what the self actually is, I will not address this objection in any depth. However, it is clear that those who deny the object nature of the self also deny many ideas that are relatively uncontroversial. It is common to address the question of self in terms of the distinction (and complementarity) between self as subject (self as I) and self as object (self as me) (Wittgenstein 1958; James 1890). The examples that Wittgenstein gives will make the distinction clearer: “My arm is broken” and “I have grown six inches” are examples of the self as object, whereas “I think it will rain” and “I have toothache” are examples of the self as subject. As these examples demonstrate, the self may appear as both an object of experience and a subject of experience (Wozniak 2018). It is the ways in which the self appears as an object of experience that are discussed here.

Understanding the self as an object is itself a metaphorical exercise because the abstract notion of self is understood in terms of an object similar to the way in which concepts like truth and justice are understood. As per conceptual metaphor theory, abstract concepts are impoverished when they are not structured by metaphors. We may not even be able to discuss what the self is without understanding it in terms of a metaphor. The notion of an object develops in infancy and childhood through bodily experience. When we conceive of an object it is possible to make a distinction between that object and its properties. Following this pattern of reasoning,it is common to distinguish between the self and its properties. Among properties, it is possible to make a further distinction between what is essential and what is adventitious. Free will or autonomy becomes a property of the self and belongingness to a community becomes another property. Accordingly, ambition and aggressiveness become the salient properties of some selves whereas friendliness and a cooperative spirit are salient in other selves. Thus, individualist and collectivist conceptions of the self represent divergent narratives that arise based on how they view the self in terms of its nature as a material object. What remains essential is the notion of a mental entity that exists by itself (though it is attached to a body) and continues to endure as the same entity over time. Without such a notion of the self it seems impossible to entertain the notion of a soul or even that of personhood – the person over and above the biological organism.

Another schema that is in operation in understanding the self is that of a container. The notion of self-fulfilment arguably comes out of understanding the self as object in terms of a container to be filled. Looking at the world or the present moment through the lens of the question, “What will I get out of this?” is a common experience in the modern world. Dissatisfaction with what we have may assume serious proportions. This dissatisfaction arises regardless of how much is available for biological survival. Even when a person can meet all of their biological needs, the mental self remains unfulfilled and looks to gain more and more from whatever it encounters.

Consider the purported cases of bicameral mind and the observations of some anthropologists regarding the uncommon ways of understanding inner life, or rather the lack of it. In a bicameral mind, an inner voice is heard but the person considers it to be from an external source. This is similar to what is called ‘thought insertion,’ which is often understood in terms of retaining the sense of ownership but not the sense of agency (Gallagher 2000). It is the sense of agency that makes us feel that we initiate an action. In the case of hearing voices or thought insertion, this is missing. At the same time, the sense of ownership, without which the person would not even be able to realize and report the presence of a voice, remains intact. This explanation has been challenged (Metzinger 2003; Lopez Silva 2017) on the grounds that in some cases, e.g. spontaneous thought and obsessive thought, a sense of agency is lacking but the sense of ownership is intact, and subjects do not consider their thoughts to derive from an external source. Accordingly, thought insertion must be understood differently from cases of spontaneous thought because the key conditions – a lack of agency and an intact sense of ownership – do not apply to both.

Lopez Silva (2017) proposes we apprehend this difference as a distinction between for-me-ness and the sense of mine-ness. In the case of thought insertion, there is for-me-ness. The subjects feel that the thoughts are given to them but they do not feel that they are the author of these thoughts. By contrast, in the case of spontaneous thought, subjects do not consider themselves as having intended to have the thoughts, but they do understand these thoughts to be of internal origin. In the case of spontaneous thought there is no alienation from thoughts whereas in purported cases of thought insertion and bicameral mind the person experiences alienation. This difference can be captured by means of the distinction between perspectival ownership and personal ownership. When a person has leg pain, she perspectivally 'owns' it in the sense that nobody else can have it, regardless of whether the cause of the pain is endogenous or exogenous. Similarly, a thought may be seen as coming from outside or inside but the person perspectivally owns it regardless of the perceived origin. But, when it is seen as of internal origin there is no alienation from it; imbued with a sense of 'mineness,' the sense of ownership becomes stronger. This aspect – mineness – is missing in cases of thought insertion. This is because mineness is related to the source of thoughts. If thoughts are conceptualized as arising from within, they are seen as part of oneself; if they are understood as imposed from outside, the relationship becomes tenuous. The former can give rise to strong, personal feeling of ownership (mineness) whereas the latter remains at the level of perspectival ownership

When the Maori consider the self to be an observer rather than director of experience, this again presents a self-narrative that offers a unique understanding of the relation a self stands in to its acts and thoughts. The prepositions 'on' and 'in' capture this fact. Consider the statements, "The guard is on duty" and "I am in a dilemma" (Green & Evans 2006). In the former case, there is some control over the situation because the guard does his duty voluntarily. But in the latter case, there is no sense of control over the situation. This difference is expressed through different metaphors and narratives and can be extended to the whole realm of life. Given certain ways of conceptualizing the self, life can be looked at as something over which the self has no control. Events unfold, and the self can only watch or undergo whatever takes place, a mere I object that is acted upon. On the other hand, the self can also be seen as encompassing the qualities of object and agent – the self is an object that can be acted upon but possesses the property of being able to cause actions.

As per a recent study, narrative selfhood can be positively altered by some psychedelic drugs that bring lasting changes (Amada et al. 2020). Individuals who have used these substances have reported feeling greater connectedness to nature and the rest of the world with the consequence that they feel increased responsibility for their behaviour. The way the self is understood in terms of its relation to external entities is thus altered by psychotropic substances. These changes may be tracked by the metaphors that are used to describe the self and its properties.. For instance, the self and the world can be seen as two completely distinct objects, comparable perhaps to pen and paper. Under these conditions there exists a thick psychological boundary between the self and the world. If the metaphor of waves and ocean is used to understand the relation between the self and the world, there is hardly any scope for such a psychological boundary to arise. Waves may appear to be unique for some time, but they are essentially 'parts' of the ocean. Between these two extremes lies a wide spectrum of views concerning the relation between the self and the world.

The upshot of the above discussion is that different ways of conceptualizing the self can be understood in terms of differences in the metaphors used to explain the self and the narratives thus engendered. Whether it be the solid, alienated, substantive self-as-agent or the self-as-object that merely observes whatever takes place, these metaphors convey a narrative that gives shape to our perceptions of reality and gives rise to abstract entities that may affect our lives in important ways.. As Yehuda Harari observes (2015), we live in two worlds: one of physical entities and the other of imagined ones. Rivers, stones, and trees are examples of the former whereas money, nations, and justice are examples of the latter. It is obvious that imagined or abstract entities play a big role in everyday life. Even concepts as important as freedom and morality belong to the list of imagined (non-concrete) entities as they do not exist prior to their being conceptualized. As Daniel Dennett observes, morality as such did not exist before the *concept* of morality came into being (Dennett 1986). But once morality became entrenched in human life it appeared to exist independently of the concept, that is, to exist objectively and independently of any human subject. The various ways in which the self is understood present a similar case. It is through conceptualization that they exist, but although they are purely abstractions, they may function as the core of a person's worldview and the fulcrum of their life.

This implies that metaphorical understanding helps in the formation of imagined entities capable of giving rise to rich narratives, which, in turn, can play a tremendous role in shaping various aspects of life. Like metaphor, narrative must be treated as belonging to the conceptual realm. Both metaphor and narrative are primarily conveyed through the medium of language; however, they are at bottom modes of conceptualization since, as per the tenets of cognitive semantics, the conceptual system is primary and vaster than what is contained in language alone. Recall that different narratives give rise to different ways of understanding the self, but not all of these understandings will be d represented explicitly in language. Instead, they may be tacitly present in our actions and approach to life. This is in consonance with the notion of the cognitive unconscious according to which many cognitive processes, including higher-order ones, take place unconsciously (Kihlstrom 1987; Wilson 2004). Anything over and above the minimal self can be treated as part of the narrative self, including metaphorical understanding and the rich inferences it engenders, whether or not a person is consciously aware of these processes. This follows if we accept that the minimal and narrative self are jointly comprehensive in accounting for the phenomenon of self.

It is important to note that the way the self is understood continues to change over time, and this can be perceived as a series of revisions in narratives. Some studies indicate that individualism is on the rise even in countries that are seen as traditionally oriented more towards collectivism (Hamamura 2012; Hamamura & Xu 2015; Santos et al. 2017). This is considered to be a global phenomenon, arising out of the spread of capitalism, a system centered around individualistic values, to almost all parts of the globe. Externally-imposed value systems may in this way determine our inner beliefs, attitudes, and identity since, as Alasdair MacIntyre noted, we are only co-authors of our own narratives (MacIntyre 1984). The dominant ideas of our time, in addition to the tradition and culture of the community we live in, undeniably shape the direction of our narratives, and it is therefore impossible to invent the whole of our narrative independently and from scratch.

Conclusion

Substantive differences in the way the self is conceptualized can be due to the use of multiple metaphors and the narratives they engender. Given the role that imagined entities play in human life, the self can be purely conceptual and yet remain significant. Conceiving of the self in this way may entail extending the notion of narrative to include unconscious and prelinguistic phenomena. t Given the increasing evidence that many cognitive processes occur at an unconscious level, this move appears to be justified, and thus allows for the possibility of complex and changing self-conceptions that are neither purely biological nor purely internal, but materially, historically, culturally, and psychologically determined.

References

Amada N, Lea T, Letheby C, Shane J (January 2020) Psychedelic Experience and the Narrative Self: An Exploratory Qualitative Study. *Journal of Consciousness Studies* 27: 6–33.

Bergen B K (2012) *Louder Than Words: The New Science of How the Mind Makes Meaning.* New York: Basic Books.

Bermudez J L (1998) *The Paradox of Self Consciousness*. Cambridge, MA: MIT Press.

Bermudez J L (2018) *The Bodily Self*. Cambridge, MA: MIT Press.

Bortolan A(2020) Affectivity and the Distinction between Minimal and Narrative Self. *Continental Philosophy Review* 53: 67–84.

Dennett D C (1989) The Origins of Selves. *Cogito* 3: 163–173.

Dennett D C (1986) Julian Jaynes's Software Archaeology. *Canadian Psychology* 27: 149–154.

Dennett D C (1992) The Self as a Center of Narrative Gravity." In *Self and Consciousness: Multiple Perspectives*,ed. F. Kessel, P. Cole, and D. Johnson. Hillsdale, NJ: Erlbaum.

Fox K C R, Christoff K (2018) Introduction: Toward an Interdisciplinary Science of Spontaneous Thought. In *The Oxford Handbook of Spontaneous Thought: Mind-Wandering, Creativity, and Dreaming*, ed. Kalina Christoff and Kieran C.R. Fox. Oxford: Oxford University Press.

Gallagher S (2000) Philosophical Conceptions of the Self: Implications for Cognitive Science. *Trends in Cognitive Sciences* 4:14–21.

Gallagher S, Zahavi D (2012) *The Phenomenological Mind,* 2nd edition. New York: Routledge.

Gibbs R W Jr, Colston H L (1995) The Cognitive Psychological Reality of Image Schemas and their Transformations. *Cognitive Linguistics* 6: 347–378.

Green M, Evans V (2006) *Cognitive Linguistics: An Introduction*. New York: Routledge.

Hamamura T (2012) Are Cultures Becoming Individualistic? A Cross-Temporal Comparison of Individualism–Collectivism in the United States and Japan. *Personality and Social Psychology Review* 16: 3–24. doi:10.1177/1088868311411587.

Hamamura T, Xu Y (2015) Changes in Chinese culture as Examined through Changes in Personal Pronoun Usage. *Journal of Cross-Cultural Psychology* 46: 930–941. doi:10.1177/0022022115592968.

Hampton R S, Varnum M E W (2017). Individualism-Collectivism. In *Encyclopedia of Personality and Individual Differences,* ed. Virgil Zeigler-Hill and Todd K. Shackelford. New York: Springer International Publishing.

Harari Y N (2015) *Sapiens: A Brief History of Mankind*. New York: Penguin Random House.

James W (1890) *The Principles of Psychology*. New York: Henry Holt and Company.

Jaynes J (1976) *The Origin of Consciousness in the Breakdown of the Bicameral Mind*. Boston, MA: Houghton Mifflin.

Johnson M (1987) *The Body in the Mind: The Bodily Basis of Meaning, Imagination, and Reason*. Chicago: University of Chicago Press.

Johnson M (2007) *The Meaning of the Body: Aesthetics of Human Understanding*. Chicago: University of Chicago Press.

Kihlstrom J F (September 1987) he Cognitive Unconscious. *Science* 237: 1445–1452.

Kracke W H (1991) Languages of Dreaming: Anthropological Approaches to the Study of Dreaming in Other Cultures. In *Dream Images: A Call to Mental Arms,* ed. Jayne Gackenbach and Anees A. Sheikh, 203–224. New York: Routledge.

Lakoff G, Johnson M (1980) *Metaphors We Live By*. Chicago: University of Chicago Press.

Lakoff G, Johnson M (1999) *Philosophy in the Flesh*. New York: Basic Books.

Lienhardt G (1961) *Divinity and Experience: The Religion of the Dinka*. Oxford: Clarendon Press.

MacIntyre A (1984) *After Virtue: A Study in Moral Theory*. Notre Dame, IN: University of Notre Dame Press.

McMahon D (2013) Divine Fury: A History of Genius. New York: Basic Books.

Metzinger T (2001) *Being No One: The Self-Model Theory of Subjectivity*. Cambridge, MA: MIT Press.

Milliere R, Metzinger T (March 2020) Radical Disruptions of Self-consciousness: Editorial Introduction. *Philosophy and the Mind Sciences* 1: 1. https://doi.org/10.33735/phimisci.2020.I.50.

Morgan L H (1877) *Ancient Society*. London: MacMillan & Company. Transcribed by Ibne Hasan. Accessed on 15-08-2020. https://www.marxists.org/reference/archive/morgan-lewis/ancient-society/index.htm

Santos H C et al.(September 2017) Global Increases in Individualism. *Psychological Science* 28: 1228–1239. https://doi.org/10.1177/0956797617700622

Schechtman M (1996) *The Constitution of Selves*. Ithaca, NY: Cornell University Press.

Seth A K (2013) Interoceptive Inference, Emotion, and the Embodied Self. *Trends in Cognitive Sciences* 17: 565–573.

Smith J (1981) Self and Experience in Maori Culture In *Indigenous Psychologies: The Anthropology of the Self,* ed. Paul Heelas and Andrew Lock, 145–160. London: Academic Press Inc.

Sorabji R (2004) The Concept of the Will from Plato to Maximus the Confessor. In *The Will and Human Action: From Antiquity to Present Day,* ed. Thomas Pink and M. W. F Stone. New York: Routledge.

Strawson G (November 2004) Against Narrativity. *Ratio* 17: 428–452.

Taylor C (1989) *Sources of the Self*. Cambridge, MA: Harvard University Press.

Tsakiris M, De Preester H (2019) *The Interoceptive Mind: From Homeostasis to Awareness*. Oxford, UK: Oxford University Press.

Wilson T D (2004) *Strangers to Ourselves: Discovering the Adaptive Unconscious*. Cambridge, MA: Harvard University Press.

Wittgenstein L (1958) *The Blue and Brown Books*. Oxford: Blackwell.

Wozniak M (September 2018) ‘I’ and ‘Me’: The Self in the Context of Consciousness. *Frontiers in Psychology* 9: 1656 https://doi.org/10.3389/fpsyg.2018.01656

HISTORICAL NARRATIVES AND UNDERSTANDING – POTENTIAL INFERENTIAL POWER

Stefan Petkov
School of Philosophy, Beijing Normal University,
Beijing, China
e-mail: yaggdrasil@yahoo.com

Abstract:

This paper defends the view that narratives that bring understanding of the past need not be exhaustively analyzable as explanatory inferences, nor as causal narratives. Instead of treating historical narrative as explanations, I argue that understanding of history can be analyzed by the general epistemic criteria of understanding. I explore one such criterion, which is of chief importance for good historical narratives: potential inferential power. As a corollary, I dispute one of the distinctive features of narratives described by some philosophers: the non-aggregativity of narrative histories. Instead, I propose that historical narratives modestly aggregate and this aggregation depends on the success of the colligatory concepts they offer.

Key words: narratives, self-conception, metaphors, minimal self, narrative self.

1. Introduction

It is considered common sense that history offers understanding of the past. However, answering how this understanding is established has been a problem with which philosophers of science and history have struggled. (For a general account of the difficulty in obtaining historical knowledge and understanding see Danto 1985.) Furthermore, some recent publications (Currie, A. & Sterelny 2017; Beatty 2017; Roth 2017) have argued that the narrative form is not exclusive to historiography but is also used in the natural sciences. According to these authors, narration is a particularly suitable strategy when unique non-repeatable events must be illuminated (Roth 2017). In order to elaborate on the problem of how the medium of historiography, i.e. the narrative form,[1] provides understanding two questions must be answered.

[1] In this text I will focus exclusively on the epistemic function of narratives to generate understanding. I do not deny that genuine historical narratives can also have aesthetic, rhetorical, or moralistic function. It is also possible (for example) that a narrative that aims to promote moral ideas can also convey epistemic information about the events and actions it narrates. Furthermore, I shall not equate epistemic narratives with explanatory narratives, for reasons that have to do with the features specific to historical narratives (see section 2). I reserve the label "explanatory narratives" only for narratives that can be easily characterized as explanations. Such explanatory narratives must meet the following conditions: they explicitly aim to answer a why question; they display a clear inferential explanatory relation; and we are able to distinguish between their explanandum and explanans. I suspect that explanatory narratives are analyzable by the existing literature on the logic of explanation, whilst historical narratives broadly should be more properly analyzable by the general criteria of understanding. Moreover the criteria of understanding should be applicable to both, given that the goal of explanation is to reach an understanding of the explanandum (Grimm 2010).

Firstly, do narratives offer a distinct epistemic strategy for reaching understanding, or are they completely reducible to either explanatory inferences or descriptions of causal dependencies? Secondly, if narratives offer a distinct strategy for reaching understanding, what are their characteristics and what could be their epistemic evaluation?

After Hempel (1942), the most appealing approach to address these problems was to equate understanding of the past with the availability of explanations for it, therefore treating narratives as explanations. Some authors have further classified narrative explanations as a type of causal explanation (Gangl 2019). For others, historical explanations boil down to causal counterfactuals (Beatty 2017). Finally some philosophers have also generalized that explanatory narratives bring understanding because they trace causal trajectories (Currie, A. & Sterelny 2017).

On the other side of the debates are philosophers who defend the view that narratives offer a unique strategy for attaining historical insight. Danto (1985) has argued that narration conceptualizes historical events and establishes relations between these events that are not necessarily causal. Recently, Roth (2017) has explored the idea that because narratives have conceptually constructive constituent parts, narration could be primary to any potential causal inferences that can be based on it. Finally, Kuukkanen (2015) has also expounded on the thesis that historical narratives have unique features that enable them to generate understanding of history. For him historical narratives contain a central argument and supporting evidence; consequently, grasping this central argument is essential for reaching understanding.

Similarly to Danto, Roth, and Kuukkanen, I also argue that narration can be a distinct source of understanding. The impossibility of completely reducing all epistemic information provided by historical narratives to either any type of inferential accounts of explanations or to causal explanations arises from two unique characteristics of historiography. Firstly, there is no rigid standard for writing history, that is, historiography is stylistically rich. This richness can be attributed to the changing epistemic and pragmatic goals of historians – goals which lead to different selections, interpretations, and descriptions of the past. Consequently, if we assume that explanations are broadly inferential then not all narratives can be taken as explanatory. This is because narration need not necessarily organize historical information into argumentative structures in order to convey a sense of understanding. Therefore, my view departs from Kuukkanen's view in positing that purely descriptive narratives can also illuminate the past.

Secondly, historical narratives also depend on conceptual colligation. Conceptual colligation in history permits the synthesis of qualitatively different sources of information and the organization of events that are spatially and temporally discontinuous in a holistic structure. All of this makes it difficult to reduce the work of historians to the mere construction of counterfactual explanations or narratives that describe a chronology of causal dependencies between past events.

Therefore, I argue that narratives can bring understanding of the past without necessarily displaying a clear explanatory structure.[2] Instead, I propose that historiography can receive epistemic warrant from the more general criteria of understanding. I focus on one of these criteria: potential inferential power. This inferential power could be accounted for by the conceptual colligation in these narratives. If successful, colligation can help historians to incorporate newly acquired historical sources into narratives they themselves enrich. Narratives provide (at least) a minimum level of understanding just by organizing their contents into a manageable epistemic structure. If this content is factive, inferences built upon it can also extend our understanding of the past. This process is demonstrated when archaeological discoveries are made based on textual sources.

As a corollary, I argue critically against one of the specific features of narratives that Roth lists: their non-aggregativity. I claim that good historical narratives at least moderately aggregate, and this aggregation depends on the success of the conceptual colligations they offer.

To give an outline of the paper: **In the second** section, I provide an overview of two of the main characteristics of historical narratives, namely, logical amorphousness and stylistic richness. I argue that narratives need not be structured into arguments with supporting evidence in order to offer insight into history. **In the third section**, I defend the claim that narratives that offer understanding can do so without either 1) being exhaustively reducible to descriptions of a chronology of causal dependencies or 2) serving as a basis for counterfactual reasoning.

[2] i.e. they do not need to contain a central argumentative inference with clearly definable explanans and explanandum.

I defend this claim by arguing that the central task of historians is to make sense of the past by writing epistemically manageable expositions. This task is achieved through conceptual colligation, which permits historians to link events in ways that might defy a typical causal order of dependence.

In the fourth section, I outline the epistemic characteristics of understanding. I also introduce the notion of potential inferential power as a form of evaluation that identifies good historical narratives. As a corollary, I claim that good historical narratives should at least moderately aggregate.

2. Narratives – logical ambiguity and stylistic richness

Assuming that explanations represent the chief vehicles of understanding, Hempel (1942) found two major obstacles to fitting historical narratives within his deductive nomological model of explanations. The explanatory function of narratives was accomplished neither by 1) invoking general laws nor by 2) neatly ordering historical information into logically robust inferences from an explanans to an explanandum. Subsequent philosophers of history did not see these peculiar features of narratives as a drawback (for an overview see Danto 1985; Kuukkanen 2015). Instead, they attributed the lack of general laws in history to its subject matter. Historians study complex, non-repeatable events, and they strive to understand their unique characteristics. In their view, even if general laws that govern history could be found, the importance of auxiliary empirical conditions would greatly outweigh the relevance of these general laws in providing explanations of unique events.

The logical ambiguity of narratives and their stylistic richness can in turn be traced to the constantly changing methodology of historians (Roth 2017). This stylistic diversity can be attributed to the fact that historians do not have a unified goal in the study of the past. Instead, their epistemic goals to organize and narrate the available historical information are constantly changing. For example, a humanistic approach to researching Colonial America could construct a narrative based on surviving documents and the interpretation of artifacts. The central aim of this narrative could be to convey the uniqueness of the daily life of people during that era. No overarching generalizations or specific argument would be necessary for the success of such a narrative (for a particular example see Kelso 1984; for a general introduction to the humanistic approach see Orser 2017). Consequently a descriptive narrative would be sufficient for the reader to understand the lifestyle of the people during the period.

On the other hand there are histories such as Virginia Smith's *Clean: A History of Personal Hygiene and Purity,*which conceptualize and offer arguments, in this case tracing the changing standards of cleanliness, purity, and hygiene. Such a project also depends on synthesizing and interpreting archaeological artifacts and documents, but because this project is conceptually oriented, arguments that justify the definitions of central notions and their role in organizing historical data are essential (Smith 2007). Such arguments may not display the logical robustness of lawful deductions as envisioned by Hempel, and yet Smith's book does offer a series of arguments about what has been considered clean, pure, and hygienic in different eras. Once readers have grasped these arguments, they can interpret them as explanations of the social and cultural factors that have modified the meaning of these concepts through time.

Recently, the view that historical arguments in narration are essential for understanding the past has been defended by Kuukkanen (2015). He has tried to downplay the expositional complexity of historiography by claiming that its stylistic amorphousness is only superficial. According to him, contemporary historiography contains two clearly distinguishable elements: 1) a central thesis given as an argument and 2) supporting evidence. He claims that historical understanding is possible even without knowledge of the specific details of events.. These facts without a doubt enrich our knowledge, but their epistemic significance is overshadowed by the importance of grasping the narrative's main argument.

Although appealing and certainly correct for some contemporary historical texts, Kuukkanen's view cannot be generalized. Beside historical projects with a humanistic bent, medieval annals and chronicles are also descriptive and do not contain explicit theses. Kuukkanen has tried to avoid these cases by denying that such narratives represent genuine historiography, as he considers proper historiography to always contain a central argument.

There are several possible criticisms of his normative position. Firstly, it is counterintuitive to suggest that past historians did not gain knowledge and understanding of the past by composing and reading chronicles. Consequently, it is also peculiar to claim that contemporary historians who focus on constructing arguments have not understood anything about the past by reading the work of past historians, as much of this work did not assume an argumentative form. Re-arranging descriptive information under new concepts and constructing new narratives that employ an argument structure may well be the best method of organizing and synthesizing our knowledge of the past. However, this makes a case for producing new (and better) accounts of history, not for disputing the epistemic value of descriptive history per se.

In addition to undermining the value of historical accounts produced in the past, Kukkanen's thesis poses a controversy for contemporary historiography. For instance, in his review of Kuukkanen's book, Froeyman (2017) has pointed out that some contemporary histories are mainly descriptive, and not argumentative; however, we still believe that these works offer understanding.

Further support for the view that understanding of the past might not always require the composition of argumentative narratives can be found in Roth's analysis (2017). According to Roth, narration is a distinct type of explanatory strategy. Due to the logical amorphousness and stylistic richness of narration a clear separation between an explanans and its explanandum is not always possible. In many situations, the narrative creates its explandum (by conceptualizing the historical sources) and then offers an explanans. Therefore, descriptive history may be an epistemically legitimate approach to reaching understanding of the past.

Thus far I have assumed by appeal to intuition that historiography offers understanding of the past and have argued that this understanding is not necessarily reached by constructing explanatory inferences or arguments. I have traced the epistemic distinctiveness of historiography, noting several features that characterize narration – the lack of general laws, the logical ambiguity, and the possible lack of explicit argumentative structure. However, I have yet to show that non-argumentative narratives cannot be analyzed as tracing causal trajectories. It could perhaps be the case that narratives are causal explanations that simply assume a descriptive form. I have also left open the question of how understanding can be achieved by descriptive history.

3. Conceptual colligation, causality, and counterfactuals

One might suggest that because history describes past events in accordance with the traces they have left behind, narratives structure inferences about causal information. Understanding of the past is gained by obtaining knowledge about causes. The stylistic richness of historiography might not permit the rigid application of a singular causal theory, but narratives nevertheless might be seen broadly as offering causal explanations (Gangl 2019; Beatty 2017). However, I argue that the expositional richness of narratives does not permit their exhaustive reduction to any of the major types of causal explanations – whether they belong to those that describe chronologies of causal dependencies or those that are based on counterfactuals.[3]

Briefly, my argument for the limited role of causal explanations is separated into two parts. Firstly, I argue that the role of conceptual colligation in organizing and reconstructing historical information often introduces links between events that lack strong causal connection. Even if a causal story is provided for all the events in such narratives, this causal story still cannot answer important purely conceptual questions about the selection and organization of these events, questions which are essential for understanding history.

Secondly, along with Carr (1965) and Evans (2004), I dismiss the usefulness of counterfactual reasoning for understanding history. This is because in many cases the counterfactual information that we are justified in inferring from historical narratives is very limited.

Beginning with the former point I must first clarify what conceptual colligation is and why it is essential for historiography. The term colligation was coined by Whewell (1840). He explored the

[3] The core analyses of causality as it relates to explanations can be typified as based on processes (Dowie 2000), mechanisms (Salmon 1984), manipulations (Woodward 2003), and counterfactuals (Lewis 1973). For brevity's sake and in acknowledgment that such a categorization does not do complete justice to the specific details of these accounts, I will classify them under two rubrics – chronologies that describe causal dependence and counterfactual accounts. The chronological dependence accounts would then be those accounts which follow a chronological causal order, as causes always precede their effects. Such accounts would describe causal dependence as a matter of interactions, processes, or mechanisms.Counterfactual accounts on the other hand capture cause/effect relations in terms of counterfactuals or manipulations.

constructive role of conceptualization in relation to scientific discovery and developed the idea that arriving at the right concept is essential for unifying previously unrelated facts within a given theory. Accordingly, both Kepler's application of the concept of elliptical orbits and the subsumption of Kepler's three laws of planetary motion under Newton's conception of acceleration are theoretical judgments, not cases of simple generalization based on objective information. The plotting of singular planetary observations into elliptical paths and the unifying of Kepler's laws of orbital motion under Newtonian mechanics were both possible only with the introduction of a subjective conceptual element – that of elliptical motion and of acceleration respectively (see Foster 2004 for a more detailed analysis).

Philosophers of history have also acknowledged the role of conceptual colligation in narration (see Kuukkanen 2015 for an overview). They have suggested that the available historical information based on documents or artifacts can be seen as the objective element in a colligation; however, in order to make sense of that information and form a narrative from it a subjective element is also needed. This requirement is fulfilled through the introduction of suitable unifying concepts. The conceptual element in writing history thus permits the synthesis of potentially temporarily discontinuous events into a unified structure. Achieving such unified structure is essential to understanding the past, as it creates epistemically manageable expositions by offering conceptual relations between the narrated events. As such, this synthesis may introduce contextually unique relations between events – relations that may go beyond describing causal dependencies. This idea can be briefly illustrated by going over two examples, one of singular narrative sentences and another of book length narratives.

The first kind of narrative may be parsed in terms of ideas that go back to Danto (1985). Any sentence that relates two distinct past events may be considered as constituting a minimal notion of narration. Given this, let's consider how the sort of sentences that we might encounter in a book of the history of physics might function. Here is one example: In 270 B.C. Aristarchus anticipated the theory which Copernicus published in A.D. 1543. Here we are not connecting one past event to another (temporally prior) event through a relationship of dependence since Copernicus' discovery is not contingent on Aristarchus nor vice versa. One could protest that such sentences serve purely rhetorical purposes and are not essential for understanding, but examples of colligations that organize whole narrative expositions are also abundant. Kuukkanen (2015) for instance analyzes Khrushchev's Thaw as such a case. The concepts of cleanliness, purity, and hygiene in Smith's book are also cases of colligatory concepts that are used for unifying information about culturally shifting hygienic habits, standards of beauty, and religious practices that are not otherwise strongly causally related to one another (Smith 2007). Such colligations could debatably be supported by an argument, but this argument could play only a justificatory function, that is, it could aim to show only that the chosen concepts provide the best or most relevant way of grasping and unifying historical information. Importantly such argumentative justification need not serve an explicit explanatory function.

An option to downplay the role of colligation would be to adopt the idea that colligation has only a limited heuristic function. It is possible that colligatory concepts are used as initial scaffolding for the presentation of past events. Whenever there are no direct causal relations, the goal of the historian would be, if possible, to discover all the missing pieces in a causal puzzle that would link all the narrated events. With the discovery of new documents and artifacts, these gaps could be progressively filled with causal information. Consequently the role of speculative concepts should be moved to the periphery or they should be discarded altogether (Gangl 2019; Currie, A. & Sterelny 2017).

Such a position might be intuitively appealing but it is bound to fail as a reductive strategy. If we accept Roth's and Danto's arguments that conceptual colligation in narration creates its subject matter, then whatever causal story supports the events described would always be secondary to the initial conceptual colligation. Naturally a causal story would always be possible; so far as narrated events are factive there will always be a set of ontic conditions upon which these events causally depend. However, no causal story – even a complete one – can replace the function of colligation and answer why certain events have been selected and why they have been described in a particular way under a particular conceptual framework.

These questions are essential to understanding history and their answers need not be exclusively causal. For instance, a historian of personal hygiene habits could study the production of toothbrushes in Western societies. Without a doubt there would be a causal story that could trace the arrival of the first toothbrush from Asia and a causal story that could account for the growing production or the innovations in constructing these artifacts. Perhaps we could even answer why these objects have ended up in the locations where archaeologists have found them. However, such causal stories would not enrich our understanding in several important ways. For instance, we could not explain why Chinese inventors created a toothbrush resembling modern ones as early as the Tang Dynasty but Western travelers did not bring them home before the fifteenth century. Addressing this problem will require first investigating 1) what is considered 'clean' in different cultures and 2) how this standard has changed over time in Europe. Once these facts have been established, it is possible to construct complex causal stories.

The second way to analyze narratives as causal explanations would be to look at the selection and organization of narrated events as underwritten by implicit claims about which events are 'difference-makers' in a counterfactual sense. Answering the question of why specific events have been incorporated into a narrative means showing that these events *mattered* – that if they had been different then the present would also have turned out to be different than it is. The view that narratives describe the events that matter, that is, those events that have shaped our contingent present, has been explored in a series of articles by Beatty (2016, 2017).

There are several reasons why the application of counterfactual reasoning to the study of history is limited. Due to a lack of surviving artifacts and other evidence or to the progressive deterioration of such evidence, there may be too many gaps in our knowledge regarding certain historical events or epochs to reason correctly about them. Conversely, there may also be such an abundance of evidence that multiple interpretations are plausible, and thus determining which is authoritative is not likely or possible (Danto 1985). There is also a simpler methodological argument for the limited use of counterfactual reasoning by historians, but in order to elaborate on it I must first provide a sketch of Beatty's thesis.

According to Beatty, a historical counterfactual should not be read in isolation. This is because the understanding a counterfactual offers derives from the fact that it is an element of a sequence of statements all of which are contingent upon the assertion of the counterfactual. For Beatty there are two ways to construct such a series. One is forward-oriented and aims to answer the question: If we take the counterfactual as a starting point, what would be the subsequent events that depend upon its assertion? An example of this kind of thinking is captured in the following question: "What would have happened if Hitler had not started Operation Barbarossa?" The other way of constructing a counterfactual is backward-oriented and aims to answer the question: If we take some temporal point subsequent to the time point of the counterfactual, what are the next events on the timeline upon which the future point is contingent? An example of this could be: "How do the events described in the 1,245 pages of the *The Rise and Fall of the Third Reich* depend on the statement, 'Throughout most of the night from Sunday to Monday Hitler paced up and down his room in the Kaiserhof hotel on the Reichskanzlerplatz, just down the street from the Chancellery' ?" (Shirer 1961, p. 21).

Whilst the first case is rightfully judged by historians as a basis for historical fiction and not for genuine history, the second according to Beatty is a basis for genuine historical reasoning, because backward contingencies show us which contingent events our present depends on. In other words, writing history is a matter of writing what has made our present the way that it is.

Regrettably, however, Beatty's argument is flawed.Both forward- and backward-tensed counterfactual statements are equally justified – but only if we possess a well-defined set of possible worlds that establish their status as true on a given timeline. Such a set would be deductively closed and allow us to judge which events are contingent (those events that are true for some subsets of possible worlds in our universe) and which events are necessary (those events that are true for all subsets of possible worlds in our universe). If such a set were given, we would also be able to know which events are contingent upon certain past events (that are on a particular timeline) and which events occur of necessity (on all possible timelines).

However historians are not clairvoyant. They have a factual, non-hypothetical characterization of only one timeline – ours. Moreover, all of the events on that timeline have equal probability, i.e. a probability of one, since all of them have already occurred. It is also unclear how we can justify the claim that a given event is contingent upon a specific past event. To clarify this point, I will use a sample from Suetonius' *De vita Caesarum* concerning the details of the assassination of Julius Caesar:

> It was just about the fifth hour of the day when he set off. As he went, someone handed him a note containing details of the plot against his life, but he merely added it to the bundle of petitions in his left hand, which he intended to read later. (Suetonius 2007, p. 47)

We can assume that both entering the senate and being killed are contingent events in Caesar's life. We may also presuppose that the successful assassination is contingent upon Caesar's not reading the letter, as these are indeed the events that occurred. Caesar did not read the letter and then he was killed. However, if we consider a counterfactual scenario in which Caesar did read the letter, we are not thereby justified in claiming that he would not have been assassinated, because there is no non-arbitrary way for us to define the possible branching timelines of which the timeline described by Suetonius would be a member. According to the only metaphysically factual account of possible worlds – that of Lewis' possible world realism (Lewis 1973) – the closest possible world would be the one where everything remains the same, save the assertion of the negation of a single factual statement. Following this reasoning, in the closest possible world, Caesar did read the letter but he still entered the senate and was killed. Naturally, there is also a close possible world where he did read the letter and was not assassinated. But which of these possible histories should we take into account?

Without possessing complete knowledge of the events in each of the possible branching timelines – events which we can then label as contingent or necessary – the problem is even thornier. For example, if one were to write a history of the Roman Empire and failed to mention the letter incident, a reader familiar with Suetonius would be uncertain whether this history described our past or not. After all, if I fail to mention the letter incident, it might be the case that I am describing the timeline where no such letter has been written or the one where Caesar reads it but still enters the senate (i.e. a timeline where the reading of the letter is immaterial to the successful assassination).

There would of course be a historical fact of the matter for our timeline, but this fact would fail to show that the successful assassination is contingent on the fact that Caesar did not read the letter since the assassination is factually asserted in both of these written histories – the one that mentions the letter and the one that does not. Determining that the assassination attempt would not have succeeded if Caesar had in fact read the letter is thus a purely arbitrary choice. After all, splitting the timelines from the point of the reading of the letter depends on a timeline which has occurred and on a purely hypothetical one for which we have zero factual information.

To conclude, I believe similarly to Carr (1965) and Evans (2004) that counterfactual reasoning has very limited use in understanding history. Furthermore, given the aforementioned role of conceptual colligation, one could address the question of why certain events are worth narrating in a given historical project and why other events are omitted without invoking messy counterfactual reasoning. Returning to the quote from *The Rise and Fall of the Third Reich*, the fact that Hitler was nervously walking about in his room between 29th and 30th of January 1933 could or could not have mattered for his election as chancellor. On the other hand, it is obvious that the sentence in the book is included for dramatic effect or to show us something about the character of Hitler, etc. Therefore, the role of conceptual colligation in organizing and selecting the events of a narrative is primary in understanding why these events are narrated at all.

This does not mean that causal information is not important for reaching an understanding of history. It is. Moreover, as we shall see, reasoning about the causal net of which narrated events are elements is relevant for extending our understanding of the past. However, this is true of cases of inferential reasoning that are pursued after a given narrative has been asserted. It is not the case that the narrative itself must be presented in the form of a causal explanation or narrative description of causal chains.

To conclude, causal explanations or tracing causal trajectories could indeed enrich our understanding of the past, but this understanding remains incomplete without a primary colligation organizing the events for which we seek these causal stories. This conceptual organization that structures events or posits relations between them is essential for achieving understanding of the past. As we shall see in the next section, one of the key features of understanding is that it establishes relations between facts.

4. Understanding and potential inferential power

So far my claims have been mostly critical – I have argued that historical narratives cannot or need not fit within any of the existing accounts of explanations, causal or inferential. Narratives need not organize events in a chronological order of causal dependence, because conceptual colligation offers other creative ways to organize our knowledge of the past. I have also shown that the other option – analyzing narratives as consisting of a central inference and supporting evidence, that is, as arguments that link an explanans with an explanandum – also fails in cases where historical texts are logically amorphous or purely descriptive.

The next step in my argument is therefore positive. I offer the criteria by which we can evaluate the understanding that can be extracted from history. The growing literature on understanding has shown that although explanations play a significant role in reaching understanding of a given subject matter, this epistemic achievement could also be secured through non-explanatory means.[4] Naturally, this claim is consistent with the account of narratives provided here and suggests that narrativity might also be one of the possible distinct strategies to reach understanding of a given subject matter. In order to defend this idea, I will first explicate what the epistemic achievement of understanding is. Then, I shall focus on the features that the sources of understanding should have in order to be successful in transmitting or facilitating this epistemic achievement. Finally, I will demonstrate that these features can also be found in good historical narratives.

Most philosophers agree that understanding is a type of epistemic achievement – an achievement that involves grasping connections, that comes in degrees, and that manifests itself through the cognitive abilities to infer, generalize, transfer, and answer why questions (Le Bihan 2017; Petkov 2020; Gurova 2017; Grimm 2010; De Regt 2017). In order to facilitate this achievement the sources of understanding should exemplify or establish connections between the facts they represent. These relations in turn should enable epistemic agents to construct further inferences about the facts under scrutiny.[5]

Applying these ideas to historiography suggests that an argumentative structure need not be exemplified explicitly in a given narrative; instead, the narrative's conceptual colligation could offer conceptual links that can be used to organize the available sources of information about the past into an exposition. The unique colligatory concepts offered by such historical expositions can then serve as the raw material for further epistemic inferences or inquiries about the past.

This idea can be cashed as the claim that historical narratives can offer understanding if they are inferentially successful. This inferential power can be manifested as:

1. The successful extension of the colligatory concepts offered by the narratives to unify newly acquired information about the past.
2. The success of the assumptions derived from these narratives in leading to the discovery of new archaeological sites, artifacts, or documents, or in offering further support in favour of some historical claim for which no consensus yet exists.

Obviously, to be potentially inferentially successful such narratives must have two characteristics, one subjective and the other objective. Firstly, they should offer colligatory concepts that appeal to the historians of the era. This appeal could be considered as presenting concepts that offer strong unification in organizing the available historical sources or interesting lines of investigation into the past. Secondly, they should colligate only factually justified or justifiable inferences about past events. In other words, a minimum requirement for narratives should be that they strive to be factive.[6]

To illustrate these ideas I will go over two cases – one concerning the discovery and interpretation of archaeological sites and artifacts, and one concerning the process of extending a colligatory concept to construct new narratives.

An example of the first case could be the famous historical and archaeological riddle: "Could the excavations on the hill of the Dardanelles be

[4] For instance, Elgin (2009) has argued that understanding can be reached via an exemplar. Dellsén (2018) has explored the idea that models can provide understanding without generating explanations. De Regt (2017) has shown that understanding in some situations can be extended by visualizations and graphics.

[5] Consequently, according to this view, explanations are a very valuable source of understanding because they can be seen as directly offering inferential blueprints for their explananda (Grimm 2010; Petkov 2015).

[6] Here, I shall not offer further arguments for the factivity of historical knowledge aside from noting the rather obvious contrast between the work of a historian and that of a historical fiction writer. The historian strives to write only about events for which he has justification to believe have actually occurred, whilst the historical fiction writer also includes fictional events as links between the events for which he has justification to believe have occurred.

identified as Troy, which Homer takes as the setting of the Iliad?" This problem has captured the attention of adventurers, treasure seekers, scholars of Classical history, and archaeologists. In the introduction of his book, *Troy and Homer: Towards a Solution of an Old Mystery*, Joachim Latacz has written that grasping this question and its likely solution are essential for a beginner in Trojan studies to understand all of the more specific problems in that field. Here, of course, I shall not attempt any argument in favor of a particular answer. Instead, I will simply assert that the Iliad has helped archaeologists to make new findings and to interpret them, focusing on the basic reasoning behind attempts to discover proof of the historical existence and specific location of Troy.

Calvert and Schliemann were the first pioneers of archaeology to seriously consider the idea that the Homeric epic was not a pure poetic fiction. They believed that it contained historical information which could help them locate the site of the ancient city of Troy. Taking some of the events narrated in the Iliad as literal historical descriptions of the past has proven fruitful and allowed Calvert and Schliemann to make important archaeological discoveries.

From our perspective, the success of this project requires arriving at new assumptions and inferences that are based on the text but are not strictly speaking intrinsic to the epic. After all, given its at least partial fictional character we cannot claim that it contains literal descriptions about causal interactions or processes. As such, these early archaeologists have probably taken the text selectively. Based on the very basic assumption thatTroy did in fact exist, archaeological discoveries have been made. Of course, causal inferences were essential for these discoveries and they still play a key role in the debates about the exact location or existential status of Troy. The discipline of archaeology, more broadly, would be impossible without processes of interpreting and dating artifacts, and both of these depend on locating the objects that may be posited within a causal context of historical events.

If we take the Iliad as the first and still perhaps the most important document that has been used for archaeological discoveries and their interpretation, and if we cannot also claim that the text provides an explicitly causal story, then we have to acknowledge that the text still provides some basic (albeit non-causal) understanding of the past. This understanding must then be based simply on the narration of events upon which archaeologists have made causal inferences that have enabled them to locate the city. Thus, a narrative can generate understanding of the past even if it is not of an explicitly causal character. Rather, in certain cases, understanding can be derived from a narrative in virtue of the assumptions that we deduce from it. For instance, such assumptions were essential in discovering the archaeological sites that are believed to be the ruins of Troy.

To summarize: in the extreme case, a narrative – even if it is partially factive and purely descriptive – can still provide a basis for understanding the past if historians and archaeologists make successful discoveries based on it.

Historical understanding can also be achieved in cases where a colligatory concept and the conceptual relations it introduces into a narrative have been used to construct a new narrative that extends our comprehension of some historical period. Due to space constraints, I will exemplify this scenario through Kuukkanen's case of Krushchev's Thaw. Kuukkanen argues that the "Thaw," which signifies the political warming of Russia towards the West during the 1950s and 1960s, is a colligatory concept that has become part of standard historical language. The concept can be interpreted to imply not only a period of political detente between the Eastern and the Western Bloc, but also a period of cultural and social synergism. During this period, new cultural values were introduced in communist Russia, but these values were also mediated and transformed by the existing values of the regime. One such case is the introduction of consumerism in Russia. In her paper "Cold War in the Kitchen: Gender and the De-Stalinization of Consumer Taste in the Soviet Union under Khrushchev," Susan Reid (2002) describes how the introduction of consumerism was shaped by contradictory values, i.e. both the Western ideal of abundance and the peculiar ideology of the regime that diminished bourgeois values, such as the accumulation of private property. Her historical analysis is possible only under the more general conceptual framework of the Thaw, because understanding this interaction between two seemingly incompatible value systems requires a more basic conceptualization of the period as a time of greater openness and interaction between the East and West.

Kuukkanen himself notes that the historical concept of the Thaw has solidified so much information that historians often take its ontological status for granted. As such, it serves to demonstrate that 1) a successful conceptual colligation can provide genuine understanding through the inferential relations it proposes and that 2) these relations are often extended in constructing new narratives based on the initial colligation.

The implications of this case study (as well as others) lead us to assert the possibility that historical narratives can at least moderately aggregate, a claim which has been denied by Roth (2017). He has argued that different narratives identify different events and therefore different causal sequences. Because of this he has denied that narratives can yield an integrated account of what happens and why (Roth 2017). Accordingly, the claim that narratives aggregate is false – firstly when it is applied to explanations in general, and secondly in the case of historical narratives specifically.

The understanding of a given explanandum is clearly enriched by a nexus of supplementary explanations (Khalifa 2017; Petkov 2020). By extension and following the arguments presented here, it is incorrect to claim that narratives that use the same conceptual framework do not aggregate. To the contrary, as the example of the Thaw indicates, a successful colligation can lead to the enrichment of the initial narrative. Therefore, we can expect that in at least some instances historical narratives that track events of the same period can aggregate and together shape a coherent body of knowledge.

This claim, however well-substantiated it appears here, remains controversial. The question of whether aggregation is possible has engendered disagreement on multiple fronts, including: the existential status of past events, appropriate interpretations of past events, and suitable concepts for organizing events into narratives. Such disputes are characteristic of any academic or scientific discipline and are not unique to history. Naturally, as studies progress, particular basic agreements about our conceptualization of the past may be possible.[7] However, we ought to resist the stronger Hegelian claim (Danto 1985) that a single complete and best-unifying conceptualization of all past and future knowledge can be achieved. (Generally speaking, this would be a case of extreme unification.) Analyses of scientific unification (Petkov 2015, 2020), for instance, have shown that the more unifying the conceptual schema employed, the less liable to falsification it is and the less understanding it can provide for its factual domain.

From the perspective of an open future, different conceptual representations of the past are not only possible but desirable because they would capture more details and perspectives of the past. Nevertheless, these various historical conceptualizations, like scientific research programs (Lakatos 1970), should at least moderately aggregate. This aggregation would be a sign of their success in providing understanding of the histories they capture in narrative form

5. Conclusion

In this short essay I have assumed what can be seen as an internalist approach to analysing historical narratives. I critique those positions that apply external criteria for explanations, such as those derived from the philosophy of science, which claim that explicitly explanatory narratives constitute the primary vehicle of understanding. According to such views, narration must be explanatory in order to illuminate the past. By contrast,I dismiss the role of explanation and focus on the central characteristics of historiography that have been discussed by philosophers and historians: the role of 1) conceptual colligation and 2) narrative-based historical assumptions in making novel inferences about the past, discovering new information about the past, or organizing the past into new extended and richer narratives. I have also proposed that these characteristics abide by the general epistemic norms of understanding, despite the fact that they are unique to the practice of understanding history. Fundamentally, the epistemic achievement of understanding requires establishing connections, a process that is perhaps best exemplified by the construction of novel inferences based on newly introduced information (Newman 2014). Historical narratives can establish unique conceptual links between events and offer new inferences about the past based on these conceptual links. Whether or not they seek to argue or explain, it is in fact clear that they play a primary role in generating genuine understanding.

[7] Exclusions are instances in which the historical information is too limited to weight in favor of a given position or interpretation. In this situation, however, we can speak of epistemic uncertainty but not of a lack of general understanding.

References

Beatty J (2016) What are narratives good for?" *Studies in History and Philosophy of Science, Part C* 58: 33–40.

Beatty J (2017) Narrative possibility and narrative explanation. *Studies in History and Philosophy of Science Part A* 62:31–41.

Carr E H (1965) *What is History?* New York: Knopf.

Currie A, Sterelny K (2017) In defence of storytelling. *Studies in History and Philosophy of Science Part A* 62:14–21.

Danto A C (1985) *Narration and Knowledge: (Including the Integral Texts of Analytic Philosophy of History).* New York: Columbia University Press.

de Regt H (2017) *Understanding Scientific Understanding*. New York: Oxford University Press.

Dellsén F (2018) Beyond Explanation: Understanding as Dependency Modelling. B*ritish Journal for the Philosophy of Science* (4):1261–1286.

Dowe P (2000) *Physical Causation*. Cambridge, MA: Cambridge University Press.

Evans R (2004) Telling it like it wasn't. *Historically Speaking* 5(4): 11–14.

Forster M (2004) The Debate between Whewell and Mill on the Nature of Scientific Induction. In *Handbook of the History of* Logic, ed. Dov M. Gabbay, John Woods & Akihiro Kanamori, 10–93. Elsevier.

Gurova L (2017) On Some Non-trivial Implications of the View that Good Explanations Increase our Understanding of Explained Phenomena. *Balkan Journal of Philosophy*, 9(1): 45–52.

Elgin C Z (2009) Exemplification, Idealization, and Scientific Understanding." In *Fictions in Science: Philosophical Essays on Modeling and Idealization*, ed. Mauricio Suárez, 77–90. New York: Routledge.

Froeyman A (2017) Review of Jouni-Matti Kuukkanen, *Postnarrativist Philosophy of Historiography*. *Journal of the Philosophy of History* 11 (1): 33–37.

Gangl G (forthcoming). Narrative Explanations: The Case for Causality. *Journal of the Philosophy of History,* 1–25,
DOI: https://doi.org/10.1163/18722636-12341425 .

Grimm S R (2010) The goal of explanation. *Studies in History and Philosophy of Science Part A* 41(4): 337–344.

Hempel C G (1942) The Function of General Laws in History. *The Journal of Philosophy* 39: 35–48.

Kelso W (1984) *Kingsmill Plantations 1619–1800: Archaeology of Country Life in Colonial Virginia.* Orlando, FL: Academic Press.

Kuukkanen J (2015) *Postnarrativist Philosophy of Historiography*. Palgrave MacMillan.

Lakatos I (1970) Falsification and the Methodology of Scientific Research Programmes. In *Criticism and the Growth of Knowledge.* Cambridge: Cambridge University Press.

Latacz J (2004) *Troy and Homer: Towards a Solution of an Old Mystery*. Oxford: Oxford University Press.

Le Bihan S (2017) Enlightening Falsehoods: A Modal View of Scientific Understanding. In *Explaining Understanding: New Perspectives from Epistemology and the Philosophy of Science*, ed. Grimm et al., 111–135. New York: Routledge.

Lewis D (1973) *Counterfactuals*. Oxford: Blackwell Publishers and Cambridge, MA: Harvard University Press.

Newman M (2014) EMU and inference: what the explanatory model of scientific understanding ignores. *European Journal for Philosophy of Science* 4(1): 55–74.

Orser C (2017) *Historical Archaeology, Third Edition*. New York: Routledge.

Petkov S (2015) Explanatory unification and conceptualization. *Synthese* 192 (11): 3695–3717.

Petkov S (2020) The degrees of understanding and the inferential component of understanding. *Organon F:* 1–34

Reid S (2002) Cold War in the Kitchen: Gender and the De-Stalinization of Consumer Taste in the Soviet Union under Khrushchev. *Slavic Review*, 61(2): 211–252.

Salmon W (1984) *Scientific Explanation and the Causal Structure of the World*. Princeton: Princeton University Press.

Smith V (2007) *Clean: A History of Personal Hygiene and Purity*. Oxford: Oxford University Press

Shirer W (1961) *The Rise and Fall of the Third Reich*. Simon & Schuster.

Suetonius (2007) *The Twelve Caesars*, tr. Robert Graves. Revised by James B. Harmondsworth: Penguin.

Roth P (2017) Essentially Narrative Explanations. *Studies in History and Philosophy of Science Part A* 62: 42–50.

Woodward J (2003) *Making Things Happen: A Theory of Causal Explanation*. New York: Oxford University Press.

Whewell (1840) *The Philosophy of the Inductive Sciences (1967 Edition)*. London: Frank Cass & Co. Ltd.

LEGAL WORLDS AND LEGAL NARRATIVES

Aaron J. Walayat*

Washington & Jefferson College
email: aaronjwalayat@gmail.com

Abstract:
More than a simple command of a sovereign, law is a form of moral communication, something that helps constitute the way we conceive of ourselves, our community, and our culture. In this essay, I argue that law is a form of "world projection," a way for human communities to use law as an aesthetic way to understand themselves. Within this legal world are narratives that present an idealized reflection of our world. Law has two functions, a reflective function, in which it mirrors the actual world and a reflexive function, in which it corrects undesirable aspects of the actual world. It is through these functions that law describes the narratives within legal relationships in order to say something real and important about those corresponding relationships in the actual world.

Key words: Law, rhetoric, legal narrative, reflective and reflexive functions of law.

> "Even children intuitively sense this law within us. Every child in the world will say, 'That's *my* toy.' That's property law. Every child will say, "But you *promised* me." That's contract law. Every child will say, 'It's not my fault. He hit me first.' That's tort law. Every child will say, too, 'Daddy said I could.' That's constitutional law. Law ultimately comes from our human nature …."
>
> – Harold J. Berman[1]

I. Introduction

In their book on legal argumentation, the late United States Supreme Court Justice Antonin Scalia and the legal writing expert Bryan A. Garner discuss the clearest example of narrative in legal practice: a Statement of Facts. In the United States, most jurisdictions require that attorneys' briefings include a Statement of Facts between the Introduction and the Analysis. A Statement of Facts is meant to tell a story, relaying the events which led to the parties' dispute.

While a Statement of Facts should not contain argumentation, Scalia and Garner recommend that it should still be written in a persuasive manner. In their discussion, Scalia and Garner highlight something important about the legal profession: Lawyers are, first and foremost, storytellers, with the first step of a legal argument being the construction of the factual background followed by argumentation on why the facts support the overall case.[2]

It is an old maxim that law provides redress for injury.[3] Legal practice, then, becomes a matter of presenting the facts of the case and explaining why the facts reveal an injury that the law was designed to redress. Law generally has few legal narratives, with the history of legal practice revealing how factual case narratives are retellings of the legal narrative that are intended to show why the client's story has been told before and what the proper ending and redress should be.

Individual events in the actual world have no legal significance by themselves and must be paired with other events to create stories actionable under

[*] Associate Attorney, Tucker Arensberg, P.C., Adjunct Faculty, Washington & Jefferson College, Judicial Law Clerk, Hon. William S. Stickman IV, 2019–2020. J.D. 2019, Emory University School of Law. B.A. 2016, Washington & Jefferson College.

[1] John Witte Jr. and Christopher Manzer, "A Prequel to Law and Revolution: A Long Lost Manuscript of Harold J. Berman Comes to Light," *Journal of Law and Religion* 29 (2014): 149–150.

[2] Antonin Scalia and Bryan A. Garner, *Making Your Case: The Art of Persuading Judges* (St. Paul: Thomson West, 2008): 93–94.

[3] William Blackstone, *Commentaries on the Laws of England*, vol. 1 (Oxford: Oxford University Press, 2016): 23.

the law. To kill a man, for example, bears little legal significance on its own. To kill a man with *malice aforethought*, however, tells a different story: The killing is now a *murder*.[4] Add an extra event however, i.e. that the killer killed the victim in self-defense, and the story changes again. Now, the law does not see the killer as a murderer, but rather as someone who killed a man in self-defense.

Legal practice draws on putting together facts to explain how the case presently before the court is like the central narrative. Facts are introduced to show that a killer killed a man. The prosecution introduces additional facts to show that he acted with *malice aforethought*, making the killer a "murderer." The defense introduces additional facts to show that the killer acted in self-defense, absolving him of murder. Both sides will bolster their interpretations of the events in the narrative by providing concrete examples. Common law systems, like the United States, would introduce case law, showing how a court (in a previous case) had found another person, similarly situated, to fit into the legal narrative and why the present case is like the previous one.

Implicit in this practice is a comparison of narratives: the narrative of the present case and the narrative implicit within the law. The case narrative is the story itself, but the point of legal argumentation is to persuade the court that the case narrative must be concluded in the same way as the implicit legal narrative. The lawyer's responsibility is to convince the court that the case narrative has been told before – to show why the current case is like a previous case and why the facts mirror the implicit facts of the legal narrative.

The central matter of the law is this: The law provides a narrative, whether gleaned from case law or inscribed in a legal code. The lawyers proceed to introduce facts to show why the present case is (or is not) a retelling of the narrative. While this is a mostly uncontroversial description of legal practice, I want to take this a step further. Legal narratives "project" an idealized (but not utopian) legal world. Consider the Pennsylvania Statute for murder:

> **§ 2502. Murder.**
> **(a) Murder of the first degree.—**A criminal homicide constitutes murder of the first degree when it is committed by an intentional killing.
> **(b) Murder of the second degree.—**A criminal homicide constitutes murder of the second degree when it is committed while defendant was engaged as a principal or an accomplice in the perpetration of a felony.
> **(c) Murder of the third degree.—**All other kinds of murder shall be murder of the third degree. Murder of the third degree is a felony of the first degree.[5]

Note the narrative inherent in "murder of the first degree." The provision implies two actors: a killer and a killed. The act is "homicide," the killing of another person. Secondly, the killing is "criminal," meaning it is in violation of the law. This is the definition of murder within the statute. Now, to become a murder in the first degree, it must be committed "intentionally." The statute therefore provides a narrative within murder of the first degree: A person killed another person (homicide), in violation of the law (criminal), with intentionality (intentional). The "narratives" in the above statute are not, admittedly, thrilling, given their lack of detail. But, they are narratives nonetheless.

The relationship between the actors creates legal narratives which are used to provide a more concrete description about how these things relate to one another in the actual world. Further, these legal narratives created in the law are not constrained to interpretation within the law, but also say something real and important about the corresponding objects of legal subjects in the actual world.

II. Legal Rhetoric

The comparison of law and narrative – indeed, the notion of law *as* narrative – is not new. There is a wealth of scholarship within the broad "law and literature" movement which seeks to understand the relationship between law and literature, including from writers like James Boyd White, Martha Nussbaum, and Stanley Fish. Further, other scholars like Judge John T. Noonan, Jr. and Harold Berman also touch upon the uses of narrative within the legal system.

James Boyd White has made a particularly elegant contribution to the study of law and literature. To White, law is most useful when interpreted as a

[4] At common law, *malice aforethought* was a necessary element for a murder. All jurisdictions in the United States now have specific requirements for murder and no longer rely on the common law for crimes like murder.

[5] Pennsylvania Statutes, Title 18 Pa.C.S.A. § 2502.

type of rhetoric, which is the "central art by which community and culture are established, maintained, and transformed."[6] Law is a form of "constitutive rhetoric," existing from the point of view of those who actually engage in its processes, as something people do and something people teach. The legal process, to White, is a rhetorical activity.[7] It is the study of constitutive rhetoric, therefore, that contributes to the way we constitute ourselves as individuals, as communities, and as cultures.[8]

John Noonan is concerned with the use of legal metaphor, though his primary concern is with law's tendency to "mask" the humanity of human beings. In the legal process, human persons are often given legal masks in an attempt to hide their shared humanity with the other human beings participating in the legal process. Noonan's legal masks refer to "a legal construct suppressing the humanity of a participant in the process,"[9] the most historically serious example being the legal mask of a person as "property." However, this concept also applies to references to the government as a "sovereign" or the current judge before the case as "the court." This literary aspect of law hides the humanity of individuals for the purposes of judicial efficiency; certain decisions, for instance, are easier to make when a participant is referred to as "the Accused" as opposed to their name.

While Noonan's book is not itself an example of the law and literature movement as a whole, it acknowledges the importance of narrative in the law. Noonan is primarily concerned with the characters of the story. While he acknowledges the existence of "roles," he warns against the application of the role as a mask. His emphasis on unmasking the humanity of the participants is an argument that the law must represent persons properly and not simply mask them as legal constructs. Thus, the argument is not that persons are not translated into the law, but that persons must be translated into the law with all of their humanity attached.

Harold Berman was a preeminent legal historian who spurred the important law and religion movement. While he acknowledged law's distinct role as a legal science, he saw law as more. Much like White would later write, Berman saw law as a "social activity" in which certain norms are formulated by legitimate authorities and actualized by persons subject to those authorities.[10] The legal process, which includes legislating, adjudicating, administering, and other conduct by legitimate officials and legal actualization, involves obeying, negotiating, litigating, and other conduct by legal subjects. Law is rules, certainly, but it is also the social and political processes of formulating, enforcing, and responding to those rules.[11]

Berman compares law to language as both are "social events" in that they bring together participants in a conversation or dialogue and also bring together all members of the language-community.[12] Functioning properly, legal language creates venues for speaking of and hearing conflicts and ways of resolving them. It creates new channels of negotiating, cooperating, and allowing a community to live together. Legal terms, therefore, "create relationships and rights among persons." Thus, these terms need formality and ceremony – words "which show the legal relationships overlaying the parties," and a setting like a courtroom to show that justice is being served with impartiality and authority.[13]

Berman was also influenced by the work of rhetorician and philosopher Kenneth Burke who viewed language as a "holistic account of human interaction with other humans and with the world."[14] Thus, for Burke, the material world is already cast through a "filter" of language when humans come to understand it with the aid of rhetorical devices, scientific viewpoints, and moral systems, all of which represent interpretive lenses for material reality that can lead to different ideas of what is important about material reality or even what material reality is.[15] The sharing and identification of these motives lead to a shared identity through shared interpretations of a common symbol. Influenced by this thinking, Berman applied Burke's thought to law, arguing that

[6] James Boyd White, "Law as Rhetoric, Rhetoric as Law: The Arts of Cultural and Communal Life," *The University of Chicago Law Review* 52, no. 3 (1985): 684. James Boyd White, *Heracles' Bow: Essays on the Rhetoric and Poetics of the Law* (Madison: University of Wisconsin Press, 1985).

[7] Ibid, 688.

[8] Ibid, 690.

[9] John T. Noonan, Jr., *Persons and Masks of the Law: Cardozo, Holmes, Jefferson, and Wythe as Makers of the Masks* (New York: Farrar, Straus and Giroux, 1976): 20.

[10] Witte and Manzer, "A Prequel to Law and Revolution," 15.

[11] Ibid, 15–16.

[12] Harold J. Berman, *Law and Language: Effective Symbols of Community* (Cambridge: Cambridge University Press, 2013): 22.

[13] Witte and Manzer, 22.

[14] Ibid, 26.

[15] Ibid.

humans, as speaking beings, produce a common nature through the dramatic and rhetorical sharing of common symbols and motives, with law an especially relevant shared identity through communication.[16]

III. Legal Aesthetics

Of the many ways that philosophers approach the relationship between law and literature, White and Berman come the closest to what I believe is the aesthetic nature of law. They both argue that law *translates* ordinary life into legal argument and vice versa.[17] They also both importantly point out the nature of law as a rhetorical way of communication and community building. Law is itself an art form intended to attribute meaning to how a community conceives of itself and how it understands its relationship to itself and to other communities. But how does it do this? I think such a goal is achieved through an aesthetic process called "world projection," where observers understand how concepts within the actual world are translated into the legal world, reflecting what the constitutive community finds real and important.

In *Works and Worlds of Art*, philosopher Nicholas Wolterstorff presents a theory of aesthetics which is deeply tied to his philosophy of action. To Wolterstorff, works of art serve as a form of "world projection."[18] He argues that artists utilize works of art as "artefacts" projecting a world, a parallel reality, that presents a reflection of the actual world. This occurs in music and studio art as much as it exists in literature and drama.[19] The projected world within the work of art is a representation of the actual world with each deeply linked to the other.

Artists attempt to project a world that is "true" to what the artist's community believes to be real and important. However, artists also project the world in their own aesthetic and artistic manner, contributing, altering, and confirming how a community considers and understands itself.[20] Thus:

> Sometimes our prizing of the world of a work of art is grounded in the fact that we are persuaded, or become persuaded, that in significant respects the world of the work is true to actuality. Sometimes it is grounded in the fact that we are persuaded, or become persuaded, that in significant respects it is *false* to actuality. But sometimes it is grounded in the fact that we don't know one way or the other. We desperately *hope* that the work shows how actuality is. Or we *fear* that it does: and strangely, the evocation of this fear is what grips us in the work.[21]

Wolterstorff's book specifically explores drama and photography, yet he argues that this theory is applicable to all mediums. Let's consider one such comical, accessible, and valid example: cartooning. The American cartoonist Bill Watterson echoes Wolterstorff's theory in his discussion of the work of fellow cartoonist Charles M. Schulz, the creator of the famous comic strip, *Peanuts*. According to Watterson, "everything about the strip is a reflection of its creator's spirit. 'Peanuts' is one of those magical strips that creates its own world. Its world is a distortion of our own, but we enter it on its terms and, in doing so, see our world more clearly."[22] The influence of the *Peanuts* comic strip comes from projecting a world which Watterson describes as a "distortion" of our own. The comparison of worlds is important in aesthetics, as a projected world says much through the ways that it is like (and unlike) the actual world. Within the world of *Peanuts*, we see what Schulz believed was real and important to the world he lived in and from this appreciation, we agree, disagree, and empathize with Schulz's view of the world represented in his work.

Like works of art, laws are also an artefact which projects a world: a legal world. In Maria Dobozy's introduction to her translation of Eike von Repgow's thirteenth-century law code, the *Sachsenspiegel* (Saxon Mirror), she writes that the volume is best described as a part of the "mirror" genre. Von Repgow, writes Dobozy, sought not only to examine the law as it existed at the time, but also to study the metaphysical and intellectual dimensions of subjects as diverse as virtue, law, and rulership. Thus, writes Dobozy, the mirror is "a device for measuring practice against the ideal concept, thereby enabling the viewer to correct flaws."[23] Legal worlds, even as a "mirror," are not simply meant to be a perfect reflec-

[16] Ibid.
[17] Ibid, 30.
[18] Nicholas Wolterstorff, *Works and Worlds of Art* (Oxford: Clarendon Press, 1980).
[19] Ibid, 358.
[20] Ibid, 361.
[21] Ibid, 365.
[22] Bill Watterson, "Drawn Into a Dark but Gentle World,"*Los Angeles Times* (Los Angeles, CA), Dec. 21, 1999. Watterson is the creator of the American comic strip *Calvin and Hobbes*.
[23] Eike von Repgow, *The Saxon Mirror: A* Sachsenspiegel *of the Fourteenth Century*, ed. Maria Dobozy (Philadelphia: University of Pennsylvania Press, 1999): 7–8.

tion of the actual world, but to be what Watterson called "distorted," as it is from these distortions that we are most moved by the world projected in the law.[24]

A legal world describes phenomena within the actual world, attributing legal significance to such actions. Within the actual world, cases which lead to legal dispute are cases occurring between people. Law serves to add legal concepts to describe the relationship between two different legal subjects. In the actual world, a man kills another man. There are two subjects to this problem: a killer and a killed. The act connecting the subjects is the act of killing. Law, however, defines this action within legal terms, creating the term "murder" to describe the action taking place between the two subjects. Law adds an additional requirement to the otherwise legally insignificant action in order to give it legal meaning: It requires a showing of intentionality.

Above, I described legal practice as comparing the narrative of the case with the implicit narrative within the law itself. It is this implicit narrative within the law, the world projected by the law, which I want to explore. The narrative within the legal world goes beyond mere description and says something normative. A legal understanding of certain events within a legal world seeks to convince observers of something real and important about the actual world. By ascribing legal significance to, say, the intentionality of a murder, law implicitly defends the proposition that there *is* a difference between intentional killings and unintentional killings in the actual world.

Law is not an abstract foundation for morality, but a form of moral communication. A society communicates its basic assumptions and values by ascribing legal significance to relationships, which is formed by the legal interpretations of those relationships by society. The work of legal translation is to translate abstract concepts within the actual world into legal concepts within the legal world. It is in the legal world that a legal system describes the relationship between legal concepts, attributing legal significance to those relationships. Thus, law first translates "individuals" and "objects" into the legal world, that is, it translates these abstract concepts into legal concepts. Law then describes the relationship between individuals and objects, one example being "property." By ascribing this relationship between legal concepts, the law attributes legal significance to the legal concepts; yet it does something more. Law also attempts to attribute meaning to the relationship between the abstract concepts upon which the legal concepts were based within the actual world. By attributing a property interest, law not only provides legal protection and an enforcement mechanism for an individual's property, it also attempts to say there is something real and important about an individual's relationship to an object, understood as one of "property." By providing legal protection for property, law implicitly claims that property exists. This is not to say that concepts like "property" or "contract" do not exist without legal enforcement, but that legal enforcement implicitly persuades people that these concepts exist. It does this by making these legal concepts fundamental assumptions that must be accepted as true in order to operate in the law at all.

A. Definitions

By "abstract concept," I mean intangible conceptions of things existing in reality. Consider the following sentence: The man owns the dog. Certainly, our intuition tells us what a "man" or a "dog" is as we can conceptualize the image of both the man and the dog. One can conceive of a "man" or a "dog" based on the word in isolation. However, our conception of "man" and "dog" is broader than what the words in isolation entail. This is because our experience of man and dog attribute more meaning than the word itself holds. A "man" does not refer to our tangible conception of a male *homo sapiens*, but to a wider idea of what a man is based on our experience of man. In similar fashion, our conception of "dog" is an abstract conception beyond a tangible *canis familiaris*, based on our experience of dogs.

Our understanding of these concepts expands or contracts depending on our experience of their relationships with other concepts. The relationship between the man and the dog provided in the sentence is one of "ownership." Thus, the sentence attributes to the man the capacity to own something and attributes to the dog the capacity to be owned. These attributions may not be immediately intuitive and, as stated above, our conceptions of the man and the dog expand based on experience of the relationships that may exist between the two objects. Therefore,

[24] Watterson's use of the world "distorted" however, comes with too many unwarranted, negative connotations. A more proper term is "augmented." The legal world is not a perfect reflection of the actual world but contains intentional changes (i.e. augmentations) existing parallel to the actual world. In comparison, we should find that these legal augmentations "overlay" the actual world.

the objects in the sentence are "abstract concepts," because while we can intuitively imagine these objects existing within the real world, there is an abstraction in the fact that our conception expands or contracts depending on what the object relates to.

It is these abstract concepts that are integrated and translated in the legal world, becoming "legal concepts." By legal concept, I mean the abstract object as it is translated within the law. The legal concept of an object in the actual world is not based on the object as it is, but on our experience of those objects. Thus, a legal conception of a "man" or a "dog" is not based on the man- or dog-as-it-exists, but as it is conceived of. Most laws begin by providing a legal definition of the terms as they are used. Let's consider the United States's legal definition of a human being. The United States federal government defines the words "person," "human being," "child," and "individual" as including "every infant member of the species *homo sapien*s who is born alive at any stage of development."[25] Laws begin by defining terms in this general way. However, in similar fashion to our limited conception of man through the word-in-isolation, the legal definition of human being is not (nor really meant to be) a complete definitional conception of a human being.

The base definition of a human being is naturally limited and not completely true to our experience of human beings. At the minimum level, the legal definition of a human legally vocalizes our most limited experience of a human being – that of a member of the species *homo sapiens*. But, it remains an incomplete view of our experience of human beings. The legal conception of a human being expands given the ways in which law delineates the relationships between human beings and other legal concepts, which are in turn based on experience. It is by defining the legal relationships between legal concepts that law expands the legal concepts themselves. Our understanding of legal concepts is expanded and enhanced by seeing them interact with other legal concepts. Thus, law's primary aim is to flesh out the relationships between legal concepts.

The legal world's relationship with the actual world is complicated. A legal world will seek to reflect the actual world. Law is created for the actual world and a legal world too incongruous with its actual counterpart is practically useless. Concepts in the legal world, which are reflections of concepts in the actual world, ascribe legal significance to the institutions and concepts within the actual world. These concepts and institutions are translated into the legal world because the community, and specifically the lawmakers, believe them to be real and important aspects of human experience which lay claim to legal significance.

But law does not only have a reflective function, for a simple reflection may not realize the values of the community that the legal world intends to enshrine. Law also acknowledges that the relationships between legal concepts may not be beneficial to society as a whole. Thus, law also attempts to play a corrective role – either to amend the legal relationship by restoring it to its natural order or to change the natural relationship between concepts in order to provide greater benefit to society as a whole through ameliorations. There are many realities in the actual world that a society may seek to redress. This is because what is "natural" is not always what is "real and important" to a community, and a legal world will reflect the actual world while also performing a corrective function that brings the actual world more in line with the community's values. Law is both *reflective* of the actual world and *reflexive* of the problematic realities translated into the legal world.

I call the corrective act of creating ameliorations in the legal world law's *reflexive* function. The reflexive function serves to express something real and important about the relationship between the abstract concepts upon which legal concepts are based. This includes everything from the hot-button political questions of the day, such as the persistence of social inequalities, to something as uncontroversial as a natural person being an individual.[26] Law says something about the actual world and it does this by creating a parallel legal world through which the residents of the actual world get a better understanding of their own world.

White and Berman both argue that one of the central aspects of law is the rhetorical function that allows us to conceive of ourselves as people, as communities, and as cultures. Here, I agree with White and Berman, yet I also think Wolterstorff adds to this proposition by providing a way in which normative claims are communicated. The act of world projection creates an operable world parallel to the actual world. The legal system translates abstract objects within the framework of this world, and thus the lawyers' rhetorical task is the persua-

[25] 1 United States Code Section 8(a).

[26] 52 Pennsylvania Code Sec. 1.8 (a)

sive art of arguing that the legal world is true to reality. It is within the legal world that observers in the actual world gain a vision of what they as people, their community, and their culture represent and of the relationships that exist between such concepts. In the next two sections, I would like to explore two examples which illustrate the reflective and reflexive functions of law as well as the ways in which these functions contribute to the conceptions that communities have of themselves.

B. The Many Meanings of Marriage: Law's Reflective Function

The institution of "marriage" is one of the most expansive legal concepts. Most of the contemporary debates about marriage, specifically in the United States, stem from its constitutive definition. Robert George, Sherif Girgis, and Ryan T. Anderson, for example, argue that a marriage is, by definition, between a man and a woman.[27] This definition of marriage was changed in the United States Supreme Court's decision in *Obergefell v. Hodges*, which required all fifty states to recognize same-sex marriages.[28] The Massachusetts Supreme Judicial Court, the highest court in the U.S. state of Massachusetts, acknowledged the complexity of marriage as an abstract concept, claiming that its decision in *Goodridge v. Department of Public Health*, which expanded the definition of marriage to include same-sex couples, was only concerned with "civil marriage," in an attempt to avoid treading on the deep-seated religious, moral, and ethical convictions underpinning the conception of marriage held by many people.[29]

Marriage means different things to different people, with different legal traditions conceptualizing marriage in many different ways. John Witte Jr., for example, notes how different religious traditions conceptualized marriage in Western law, with marriage classified as a sacrament in the Roman Catholic tradition, as a social estate in the Lutheran Reformation tradition, as a covenant in the Calvinist tradition, as commonwealth in the Anglican tradition, and as a contract in the Enlightenment tradition.[30] At first glance, one would assume this points to the fact that marriage has changed in form throughout the development of these traditions, but this is not the case. Marriage has not transformed from sacrament to contract, but has rather been augmented through the development of new traditions.

The constitutive debate ignores the relationship of marriage as a legal concept with other legal concepts. The reflection of "marriage" in the legal world is not limited to its definition but rather shaped by the fact that marriage *itself* is translated and recognized in a legal world. The phenomenon of "marriage" in the legal world is therefore acknowledged to be a real and important aspect of human experience warranting lawmakers' decision to provide for it within the legal world.

Thus, the debate between no-fault divorce laws and covenant marriage laws centers on the legal concept of divorce.[31] While there appears to be friction between what can be interpreted as the economic partnership model of marriage and the covenant model of marriage, it must be noted that this clash of traditions exists within the context of divorce law. But, marriage is not limited to divorce law, and the discussion of legal marriage should not be limited to the constitutive debate.

In 2004, the United States General Accounting Office identified 1,049 federal statutory provisions in the United States Code in which marital status is a factor.[32] Naturally, the constitutive debate is central in determining the applicability of these provisions to people, but it is due to the relationship between marriage-as-a-legal-concept and these other federal (and state) statutory provisions that the law's fuller definition of marriage takes shape.

Consider evidence law. In the United States, there are two important privileges in evidence law revolving around a couple's marital status: spousal privilege and confidential marital communications. Spousal privilege provides that the spouse of an accused defendant may not be compelled to testify against their spouse. There is a division between jurisdictions about who holds the privilege, with the

[27] Robert George, Sherif Girgis, and Ryan T. Anderson, "What is Marriage?" *Harvard Journal of Law and Public Policy* 34, no. 1 (2010).
[28] *Obergefell v. Hodges*, 576 U.S. 644 (2015).
[29] Joel Nichols, "Marriage: Civil, Religious, Contractual, and More" *Family Court Review* 50, no. 2 (2012): 222.
[30] John Witte Jr., *From Sacrament to Contract: Marriage, Religion, and Law in the Western Tradition* (Louisville: Westminster John Knox Press, 1997).
[31] Amy L. Stewart, "Covenant Marriage: Legislating Family Values," *Indiana Law Review* 32, no. 2 (1999): 514.
[32] "Defense of Marriage Act: Update to 1997 Report – 2004," U.S. General Accounting Office. https://www.gao.gov/new.items/d04353r.pdf

majority of states providing that the privilege is held by the testifying spouse and may be waived . A minority of jurisdictions actually allows the defendant to hold the privilege to prevent their spouse from testifying against them. Further, "confidential marital communications" allow a person to prevent their spouse (and even former spouse) from testifying about certain statements made "in reliance on the sanctity of marriage." In this case, both spouses hold the privilege and even hold the privilege after marriage as long as the statements were made during marriage. Of course, these are general rules that differ from state-to-state in the United States and there are many exceptions to this general rule.

Milton Regan has argued that these privileges recognize a narrative of marriage which includes themes such as the importance of loyalty and trust, responsibility and identity.[33] In the context of evidence law, the provisions for marital privileges reflect something that the community sees as real and important in marriage within the actual world. There is a recognition that spouses must not be compelled to testify against each other and that there is something valuable about being able to protect statements made to one's spouse within the trust of marriage. Naturally, these rationales are controversial, and there has been a concerted effort to change these privileges based on the disparate gender impact of these privileges. Most U.S. states, however, have kept the privilege with exceptions, and in 2019only the U.S. state of New Mexico had abolished spousal privilege outright on the basis that it was misogynistic in practice.[34]

Also consider trust laws. A trust is a three-party fiduciary relationship in which a testator places property in the control of another party (a trustee) for the benefit of a third party (the beneficiary). In some situations, a trust can restrain a beneficiary's ability to assign the right of the trust to someone else. This is called a "spendthrift trust," and it seeks to provide for the beneficiary and to protect the beneficiary from their own bad economic decisions. A spendthrift trust could even prevent the involuntary transfer of the trust's interest. Thus, a creditor of the beneficiary could not take the beneficiary's interest in the trust, but would collect the part of the disbursement of the trust only after it is given to the beneficiary. In some states where a spendthrift provision is generally valid, however, the claims of a child or a spouse for alimony or support or the claims of someone who provides the beneficiary with "necessaries" may override the power of the spendthrift trust to collect the interest.[35] There is another narrative within this provision. A spendthrift trust serves as a way for a testator to provide for a beneficiary with little interference. Yet, it is not absolute, and the law sees certain relationships as important enough to violate the testator's intent. Among these is spousal support. Thus, there is an inherent narrative that the value of a marriage (even a former marriage) is powerful enough to negate the testator's intent by nullifying an otherwise valid spendthrift clause.

Indeed, law provides differing narratives concerning a marriage's relationship to property. In a community property state, the money one spouse earns in a marriage is owned by both spouses. In a common law state, if one spouse earns money, it does not automatically become the property of the other spouse. There are exceptions to these general approaches to community property, such as property brought to the marriage and inheritances. Also, the distribution of marital property differs by state. Pennsylvania, for example, is a common law state that abides by the principle of an "equitable distribution" of property. Community property states can also abide by the equitable distribution principle, though a U.S. state like California splits community property in half between the spouses.

A community property system creates a narrative. The marriage exists outside of the couple, with each individual member of the marriage claiming a property interest in the marital community. The law draws a distinction between the marriage and the participants in the marriage, with the participants possessing a property interest in the marriage-as-community. In a common law system, the difference between spouses is maintained and the marriage exists *as* the partnership between the spouses. The common law system maintains that the spouses own property as marital partners while in the community property system, the marriage itself owns the property with the individual spouses owning an interest in the marriage.

A society's valuation of marriage is connected to the way marriage relates to other areas of law. Law recognizes spousal loyalty and fidelity in the context of evidence law. It recognizes the special

[33] Milton C. Regan, Jr. , "Spousal Privilege and the Meanings of Marriage," *Virginia Law Review* 81, no. 8 (1995): 2119.
[34] *New Mexico v. Gutierrez*, No. S-1-SC-36394 (2019).

[35] Paul G. Haskell, "Teaching Moral Analysis in Law School," *Notre Dame Law Review* (1990): 1047–1048.

place of marriage in its capacity to override a spendthrift trust. Indeed, some states even envision a marriage as having an independent existence separate from its participants for the purposes of marital property (i.e. in a community property system), while others maintain spousal partnership. A community's view of marriage is not limited to a constitutive definition as the definition of marriage is elaborated through its relations to other legal concepts.

C. The Competing Narratives of Hearsay: Law's Reflexive Function

Wolterstorff has argued that works of art project a world that is often reflective of what the artist's society finds meaningful and important. Similarly, law also projects a world. In this case, the Rules of Evidence project a world. Within this legal world, the abstract concepts within the real world (i.e. statements, hearsay) are translated into the legal world, constrained by legal conceptions. The clearest example of translating abstract concepts through the constraints of legal concepts comes in the form of defining the concept itself. Thus, the law begins the translation of hearsay by beginning with the definition of hearsay. Once defined, the law establishes the concept's relationship with other concepts. In this case, the concept of hearsay is most entwined with the concept of "admissibility," that is, whether something (i.e. statements, objects, etc.) can become "evidence" to be used within the legal procedure. Law therefore incorporates these abstract concepts into the legal world, transforms them into legal concepts by defining them, and explains the relationship between legal concepts within the legal world.

The use of the word "admissible" is interesting in its own right, indicating a situation in which statements and objects are "admitted," or "entered into," evidence. This implies that these statements must "meet the threshold" in order to be considered evidence at all. One can imagine a house separated from the outside world by two gates. The first gate is "relevance," and a statement or object must first be relevant in order to pass through this gate. Secondly, the statement or object must be "admissible" – that is, free from privilege or any other rule, like the hearsay rule. Only those statements and objects that are admissible may pass through the second gate and become "evidence."

Let us consider the example of hearsay in the evidence law of the legal systems that I am most familiar with: the United States and the U.S. state of Pennsylvania. I compare the Federal Rules of Evidence, the rules that apply to the federal courts, with the Pennsylvania Rules of Evidence because they, interestingly, tell two separate legal narratives. Hearsay is an "out of court statement offered as proof of the truth of the matter asserted."[36] Generally, hearsay is inadmissible as evidence in a trial, though there are exceptions. The Federal Rules and the Pennsylvania Rules differ in one major respect. The Federal Rules deem certain statements as "nonhearsay." These include the prior inconsistent statement "made under penalty of perjury" at a trial, hearing, or other proceeding, prior consistent statements, and prior statements of identification.[37] The Federal Rules also contain a variety of hearsay exceptions, including dying declarations, public records, and many others. Certain statements made that would normally be characterized as hearsay are determined to not be hearsay for that purpose. The Pennsylvania Rules, on the other hand, do not recognize any statement as "nonhearsay." Each of the statements characterized as "nonhearsay" in the Federal Rules are provided for as "hearsay exceptions" in the Pennsylvania Rules.

Note that in the trial setting, the result would most likely be the same. A lawyer who raised "nonhearsay" under the Federal Rules would simply raise a "hearsay exception" under the Pennsylvania Rules. Nevertheless, each set of Rules provides varying narratives within the law. Under the Pennsylvania Rules, hearsay is recognized as a general rule. The exceptions specify instances where statements that would otherwise be hearsay should be accepted as justified despite their characterization as hearsay. The Pennsylvania Rules recognize this as well, providing hearsay exceptions. However, the Federal Rules take a step further. Certain statements are not simply justifications for hearsay but are so fundamentally tied to the trial process that they should not even be considered hearsay.

The Rules acknowledge that the prior inconsistent statements of a declarant-witness "made under penalty of perjury" at a trial, hearing, or other proceeding or in a deposition are admissible as both substantive evidence and to impeach the witness. A prior consistent statement is allowed if it is used to rebut an express or implied charge that the declarant is now making but only if the declarant had reason to fabricate it or if an improper influence or motive arose. An out-of-court "identification" of a person is

[36] Federal Rules of Evidence 801(c)
[37] FRE 801(d)(1).

admissible as nonhearsay. By labeling these types of statements as "nonhearsay" the law sees something special about them – there is something about them by their very nature that the public would not otherwise object to as hearsay.

I think there is something special about the fact that the above statements are being made against declarants themselves. Thus, a court will not characterize as hearsay an out-of-court statement if it is made under penalty of perjury and it is being used against the declarant-witness, either to impeach their credibility or to undermine the general case. Thus, the law seeks to say something special about these kinds of statements: (1) That the "penalty of perjury" provided by the court's supervision of the statement shields the statement from the dangerous influences contained in a statement not made under penalty of perjury, and (2) that the inconsistent statement should not be hearsay if it is being used to counteract the very person who made the statement and who now seeks to testify against the statement. The position established by the Federal Rules is not just that this is an acceptable justification for admitting hearsay evidence, but that the statement is, by virtue of these two factors, not hearsay at all.

The evidence law surrounding the admissibility of hearsay, therefore, shows both the reflective and reflexive functions of the law. Reflectively, law provides a general rule: It excludes hearsay, acknowledging a societal assumption that such a statement would not be reliable. Both the Federal and the Pennsylvania Rules, however, provide that sometimes a hearsay statement is necessary and reliable enough to be admitted into evidence. The world projected by the Pennsylvania Rules is one of exception – that all hearsay is hearsay, but there are justifiable uses of such statements. The Federal Rules's world is similar but accepts certain statements as so reliable that they should not even be considered hearsay at all.

Thus, the reflexive aspect of the law can go so far as to say something is true about a legal concept that is simply not so of the corresponding abstract concept, and thus, that there is something real and important about a legal concept (based on our contextual experience of the concept) that warrants rejecting the identitive property. Take, for example, the amusing case of tomatoes in the United States Supreme Court case *Nix v. Hedden*.[38] This 1893 case concerned the classification of a tomato as a "fruit" or a "vegetable" under the Tariff Act of 1883. The Act would tax imported vegetables, but not fruit. The case surrounded an issue of imported tomatoes. The tax collector sued the importer for the taxes on imported tomatoes, but the company responded that they should not be taxed because a tomato is, botanically, a fruit and not a vegetable. The Supreme Court decided that despite the tomato's botanical definition as a fruit, it would be legally defined as a vegetable because tomatoes are most often "used" as a vegetable, rather than being eating alone as a fruit. Justice Horace Gray, the author of the opinion, even cited another case, *Robertson v. Salomon*, in which the Supreme Court had defined beans as a vegetable rather than a seed, because even though a bean is botanically a seed, it is a "vegetable" in "common parlance."[39]

Why would the Supreme Court feel empowered to define something outside of its actual definition? A tomato, again, is an abstract concept. It is a (botanically speaking) fruit that is often used "as a vegetable" in cooking. Yet, given the tomato's relationship with other abstract concepts, the conception of the tomato changes. At first glance, it would seem that the Supreme Court is most concerned with the legal concept of a tomato's relationship with the legal concept of cooking. However, the Supreme Court was instead more concerned with the legal concept of a tomato's relationship with the Tariff Act of 1883. While the Court discussed the common parlance of a tomato in everyday cooking, it was actually looking at the common parlance of a tomato in everyday cooking for the purposes of taxation under the Tariff Act. The Court is not concerned with whether a tomato is used as a fruit or a vegetable in ordinary people's cooking for the sake of defining how a tomato is used in people's cooking, but whether a tomato must be taxed like other imported produce used in similar fashion. Thus, for the purposes of taxing people's produce, the use of the legal concept of a tomato was so much like a vegetable that the law decided that the tomato *is* a vegetable. Changing the very definition of a legal concept says something real and important about the relationship between tomatoes and the Tariff Act. Law's reflexive function is a corrective function, but "correction" is not only adjusting the legal world to represent the actual world, but also changing the legal world so that it is different from the actual world but truer to the community's values.

[38] *Nix v. Hedden*, 149 U.S. 304 (1893).

[39] Ibid. *Robertson v. Salomon*, 130 U.S. 412 (1889).

IV. Conclusion

As Berman says in the opening quote of this essay, even children intuitively make appeal to concepts of "law." It is those abstract concepts – "that's *my* toy," "you *promised* me," "it's not my *fault*" – that law seeks to translate and invest with legal significance, thus serving a communicative and pedagogical function for observers in the actual world by claiming such concepts are real and important. It is within law that these narratives are constructed both reflectively, by incorporating those things we value and take for granted in the actual world, or reflexively, by transforming undesirable states of affairs into something that we do value.

Further, it is important to see the relational value of law, with our legal conception of concepts expanding depending on the relationship that those legal concepts have with other legal concepts. It is law's interpretation of those relationships that contributes something meaningful about how we understand the corresponding relationships in the actual world. Even Richard Posner, for all of his criticism of the law and literature movement, falls into constitutive rhetoric in his economic analysis of the law.[40] By interpreting law in the language of economics, he implicitly claims that economics has something real and important to say about the way these legal relationships reflect the actual world. Indeed, given law's communicative and pedagogical function, the use of economic analysis even communicates to persons that such behavior is the proper way to act within the legal world. By interpreting law economically, observers may be nudged towards treating their actual relationships according to the economic interpretation of its corresponding legal relationships, an ironically rhetorical act.

Law and the legal narratives within the law do not exist ontologically in some parallel universe, but in the meaning gleaned from the aesthetic comparison of the projected legal world with our actual world. Law is fundamentally concerned with persons, and legal world projection enters our lives through contemplative action. The construction and projection of worlds of law exist as a communicative means of determining how persons, communities, and cultures conceive of and constitute themselves. In the words of John Noonan:

> No person itself, the law lives in persons. Rules of law are formed by human beings to shape the attitude and conduct of human beings and applied by human beings to human beings. The human beings are persons. The rules are communications uttered, comprehended, and responded to by persons. They affect attitude and conduct as communications from persons to persons. They exist as rules – not as words on paper – in the minds of persons.[41]

[40] Richard A. Posner, *Law and Literature*, 3rd ed. (Cambridge, MA: Harvard University Press, 2009).

[41] Noonan, *Persons and Masks of the Law*, 4.

References

Berman H J (2013) *Law and Language: Effective Symbols of Community*. Cambridge: Cambridge University Press.

Blackstone W (2016) *Commentaries on the Laws of England*, vol. 1. Oxford: Oxford University Press.

George R, Girgis S, Anderson R T (2010) What is Marriage? *Harvard Journal of Law and Public Policy* 34(1): 245–287.

Haskell P G (1990) Teaching Moral Analysis in Law School. *Notre Dame Law Review* 66(4): 1025–1051.

Noonan J T Jr. (1976) *Persons and Masks of the Law: Cardozo, Holmes, Jefferson, and Wythe as Makers of the Masks*. New York: Farrar, Straus and Giroux.

Posner R A (2009) *Law and Literature*, 3rd ed. Cambridge: MA: Harvard University Press.

Regan M C Jr. (1995) Spousal Privilege and the Meanings of Marriage. *Virginia Law Review* 81(8): 2045–2156.

Scalia A, Garner B A (2008) *Making Your Case: The Art of Persuading Judges*. St. Paul: Thomson West.

Stewart A L (1999) Covenant Marriage: Legislating Family Values. *Indiana Law Review* 32(2): 509–536.

Von Repgow E (1999) Translated by Maria Dobozy. *The Saxon Mirror: A Sachsenspiegel of the Fourteenth Century*, ed. Maria Dobozy. Philadelphia: University of Pennsylvania Press.

Watterson B (1999) Drawn Into a Dark but Gentle World. *Los Angeles Times* (Los Angeles, CA), Dec. 21.

White J B (1985) *Heracles' Bow: Essays on the Rhetoric and Poetics of the Law*. Madison: University of Wisconsin Press.

White J B (1985) Law as Rhetoric, Rhetoric as Law: The Arts of Cultural and Communal Life. *The University of Chicago Law Review* 52(3): 684–702.

Witte J Jr. (1997) *From Sacrament to Contract: Marriage, Religion, and Law in the Western Tradition*. Louisville: Westminster John Knox Press.

Witte J Jr., Manzer C (2014) A Prequel to Law and Revolution: A Long Lost Manuscript of Harold J. Berman Comes to Light. *Journal of Law and Religion* 29(1): 142–169.

Wolterstorff N (1980) *Works and Worlds of Art*. Oxford: Clarendon Press.

REASSESSING PERSONALITY AND NARRATIVES IN THE BRAIN AND BEHAVIORAL HEALTH SCIENCES

Paola Hernández-Chávez
Universidad Autónoma Metropolitana-Azcapotzalco & Universidad Juárez del Estado de Durango, México
e-mail: hcpaola@gmail.com

Oscar Lozano-Carrillo
Universidad Autónoma Metropolitana-Azcapotzalco, México

Abstract:
Narratives play an important role in the conceptualizations and classifications of mental disorders and cognitive dysfunctions. They recur in psychiatry, psychology, cognitive sciences, impairments' therapeutics, etc. Despite their relevance, first-person reporting and specialists' recounting of clinical cases have been understated in the literature. This is intriguing since narratives can potentially influence diagnostic statements, procedures, and prescriptions of rehabilitation treatments. They can also account for the extent to which certain disorders are normalized or pathologized within specific cultural contexts. Nonetheless, a narrative/story/description cannot be substituted for the contributions of the brain and behavioral health sciences. In Section I, we summarize three reasons that could explain the deflationary view of narratives in the clinical and neuroscientific literature: a) The brain and behavioral health sciences' aspiration to emulate successful disciplines centered on pathogen-causal models; b) The bioinspired explanatory patterns; and c) The brain and behavioral health sciences' neglect of the big picture, i.e., the interaction of components when a cognitive/psychiatric/psychological problem presents. A concomitant core problem is presented in Section II: Psychiatry's out-of-date conception of personality assumes that personality traits are fixed features of a subject's identity and that identity is a static closed system. In Section III, we challenge this view and urge brain and behavioral health sciences professionals to update their notion of personality and narrative. We conclude by offering some criteria that distinguish genuine narratives from story-like accounts (i.e., genuine narratives must be consistent, explanatory, coherent, and constant).

Key words: personality, psychiatry, narratives, mental disorders, cognitive dysfunctions.

I. Why have the brain and behavioral health sciences underestimated the relevance of Narratives?

The ***'***brain and behavioral health sciences' are the chosen term to encompass intersecting professional disciplines that share concerns related to: psychiatric classification, identification of functions, atypical psychological behaviors, theories of the human mind, cognitive impairments, cognitive dysfunctions typing, rehabilitation therapeutics, etc. Importantly, these disciplines rely on the testimony of patients and specialists; yet,, in addressing these concerns, the relevance of narrative has been overlooked due to three main reasons:

a) The primacy of causality models and assumptions, i.e., the brain and behavioral health sciences have attempted to emulate successful disciplines; therefore, they follow pathogen-causal models.

b) The focus on bioinspired explanatory patterns in the brain and behavioral health sciences.
c) The dismissal of holistic thinking, i.e., the brain and behavioral health sciences neglect the big picture, namely, the integration and interaction of different components when a cognitive/psychiatric/psy-chological case presents.

Let us analyze each of them in the following sections.

a) The brain and behavioral health sciences' causality models and assumptions

Attributing functions and purpose to organs, parts, systems, is an ancient practice (Allen & Neal 2020). For centuries, the imposition of a notion of function has played a significant role in evolutionary biology, genetics, medicine, ethology, and recently in psychology. We routinely explain how and why an organism is the way it is by citing the function of its parts (Ariew et al. 2020: 1). As Perlman (2002: 264) claims, teleology, that is, explaining phenomena in reference to their having a function, compels us to approach scientific inquiry by asking what the functions of the processes, organs, genes, behavior, etc., that we observe, are. Therefore, teleological explanations occur in many sciences. Notwithstanding the potential criticisms of this, there seems to be a consensus regarding the ineliminability of teleology, which, for some authors (Ariew 2002: 30), as long as it is restrained, it is not so bad.

The interesting question here is whether there are good alternative ways to ground functional explanations beyond ascribing "proper" functions and dysfunctions to every organ of the body, or to the brain, or to cognition in general. This question, in addition, leads us to wonder about the foundations and theory underlying diagnosis and classification systems in the brain and behavioral health sciences, as well as the common assumptions about causal dynamics that may entail reductionism.

Reducing complex phenomena to their very basic and smallest components is typically considered to be a reasonable starting point in scientific investigations. Based on this initial analysis, inferring the organizational patterns, underlying forces, and general relationships in terms of theories, laws, and ideal models is an equally suitable next step.

Reduction, understood as the simplification of complexity, is a handy tool in biology, medicine, and related sciences studying characteristically complex systems. Reduction as a guiding methodological assumption can help researchers isolate the main drivers of a disease where there may be a wide variety of causes but only a few causal pathways that are common to most cases (Andersen 2020: 80-1). From the methodological perspective, as Andersen (2020) also claims, reductionism is useful for guiding interventions: Targeting dysfunction earlier in a causal chain of cascading problems is optimal for preventing or treating disease since it is easier to intervene in less complicated or entrenched situations than at later stages when problems may have multiplied or become deeply ingrained.

In the brain and behavioral health sciences, cleavages between theory and methodology are even more frequent than in the natural sciences. One the one hand, a theory may explain a problem while failing to provide insight that guides effective intervention strategies; on the other hand, an approach may prescribe effective interventions while failing to explain the health issue at stake. This is directly related to the integration problem: When, for the sake of understanding and diagnosis, we reduce a health issue to its parts, we lose sight of the integrated whole, which is essentially to planning successful interventions, while conversely, when we fail to analyze the whole in terms of its myriad functions, we assume a target for intervention or locus of a problem that is not the primary causal driver although it is pragmatically accessible to us (Andersen 2020: 82).

In the brain and behavioral health sciences, the high-level explanatory theory oftentimes does not match with the low-level intervention practices. These are instances of what Schaffner (2006, 2011) has labeled "creeping reductions", which are the opposite of "sweeping reductions," i.e., fractionated reductions of phenomena at different levels such that the higher level is not eliminated but rather connected to the lower level. Therefore, sweeping reductions are successful and also explanatory.

In contrast, in "creeping reductions," there is an ontological disruption since the higher levels do not contain the lower ones nor are they dependent on the secondary levels in a metaphysically strong sense. They therefore lack the unificatory power of sweeping reductions. Nonetheless, even when fragmentary and patchy, creeping reductions provide useful explanations that may be very important in clinical practice (Schaffner 2011: 139).

For medical and therapeutic purposes, reductions with such limited applications are nonetheless advancing. For practitioners routinely doing the hard work of adapting theoretical intervention strategies, the key question that arises is: "Does the theoretical classification and labeling of dysfunctions and illnesses need to match with practice?" The obvious answer is: Maybe theory and practice *should* match, but in everyday clinical experience, normativity may take a back seat to pressing pragmatic issues. a The holy grail of Unity of Science that Carnap (1934) and the logical positivists so busily pursued, where the reduced parts should ontologically depend, reassemble, and correspond to the general phenomena or ideal end state, collapses when dealing with brain and behavioral health issues. In these fields, perfect theoretical reductions do not take precedence. On the contrary, sufficient explanations guiding effective interventions and successful therapeutic and rehabilitation procedures are top priorities, even when theory remains fragmentary or removed from clinical practice.

To summarize: The brain and behavioral health sciences have to survive and thrive with a shortage of ontological, metaphysical, and reductionist practices since effective intervention is their primary focus. The overarching goal is a pragmatic one: to achieve restoration and maintenance of the functional psychiatric, psychological, or cognitive system.

b) The bioinspired explanatory patterns in the brain and behavioral health sciences

For good reasons, it is not an exaggeration to claim that the brain and behavioral health sciences, in general, are still mired in notions of disease and malfunction derived from nineteenth-century biology. A commitment to chasing the pathogen agent (in terms of a bacterium, virus, gene, mutation, etc.) is an efficient tool from a methodological perspective. Here, a pattern of atomistic logic prevails: A single agent 'S' causes infection 'I.'

Pathogen-atomistic-based explanatory models in biology and medicine (i.e. agent 'S' causes infection 'I') enjoyed a glorious era in the nineteenth century. Illnesses such as cholera offer great examples of successful causal pathogen identification: bacterium vibrio cholera causes cholera. For an example taken from the last century: treponema pallidum causes syphilis. In this century: the H1N1 virus causes swine flu; the Covid-19 virus causes coronavirus disease. Along this line of thought, once the pathogen is identified, it is possible to track its dynamics and predict its course of action so that subsequent prognoses and models of intervention arise. Similar bioinspired ontologies and causal explanatory practices have persisted up to the present day in cognitive brain sciences (Hernández-Chávez 2019).

The logic behind classic pathogen-causal biology proved to be quite successful for over three centuries. But, for several reasons, it became indisputable that the reductionist paradigm in biology was not the right way to start in order to account for a vast number of physical illnesses. This conclusion can be extended for many cases that present in the brain and behavioral health sciences.

A shift occurred when some biologists stopped focusing on causally isolated entities like individual genes. As some authors noticed, atomistic explanations such as the concept of an 'organism' vanished from mainstream literature as a fundamental explanatory concept in biology (Webster & Goodwin 1982; Laubichler 2000; Huneman & Wolfe 2010; Cornish-Bowden 2006; Nicholson 2014). The retreat from atomistic thinking was accomplished once biologists embraced systematic, enriched views. Thus, contemporary biologists focus not only on genes or other isolated specimens but on much broader systems. They take into account that genes are expressed within an organism, that organisms form complex systems, that they are the agents of evolution but not the only ones, that phenotypic plasticity operates in nature (i.e., that genotypes generate different phenotypes depending on the environmental circumstances), and that the heterogeneity of organisms provides greater dynamicity and variability.

We will close this section with a warning about the dangers of adopting atomistic pathogen unicausal assumptions inspired by nineteenth-century biology for current purposes. Despite a history of successful applications in medicine, atomism is extremely limited in its capacity to either provide explanation or to guide interventions in human psychology and behavior since both of these are conditioned by biological and experiential factors that cannot be isolated and analyzed independently.. In addition to the sheer complexity of the human mind and behavioural expressions, it is clear that the conditions of human life are very different from the laboratory settings in which atomistic research, such as single-molecule DNA mapping, can be conducted most effectively. Instead of pursuing investigations into the human mind based on this traditional model, it may be preferable that innovative explanatory patterns of systems biology penetrate reciprocally in the brain and behavioral health sciences. A systems approach in these sciences would imply modeling behavior, brain functions, cognitive dysfunctions, etc., within a complex functioning framework. It would also include integration, regulation, innovation, and organization of living matter understood as a complex whole rather than a hierarchical assembly of necessary genetic components.

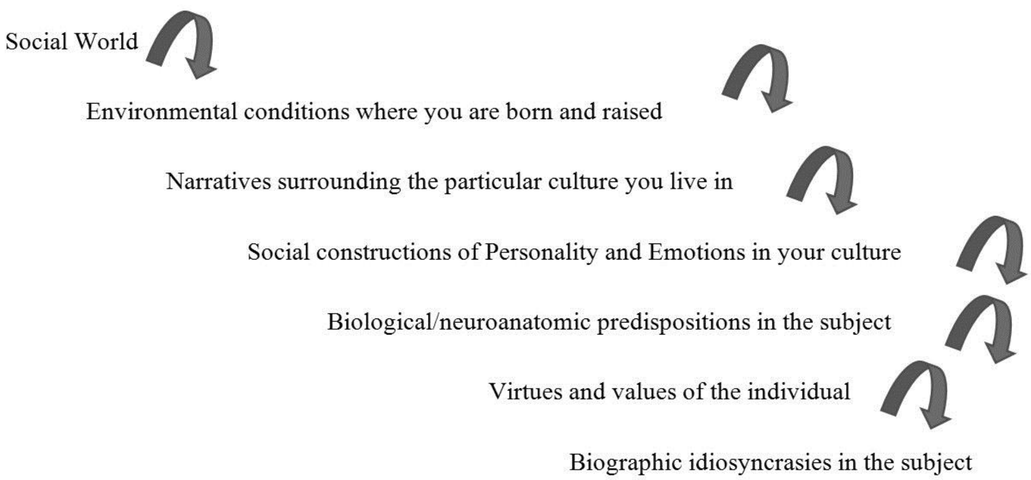
As far as the brain and behavioral health sciences go, some fundamental guidelines would be useful in forming a checklist of options to ponder. When an emergent dysfunction, illness, cognitive deficit, disruption, etc., demands an explanation (whose main purpose is to guide an intervention to alleviate symptoms rather than to yield theoretical insights), it would be essential to remember that the health issue can be the result of either: 1) a monogenic origin, such as a failure in the transcription of DNA or RNA in the formation of proteins or a mutation in the mechanisms that control for the genetic susceptibility to certain conditions; or 2) a multigenic origin, such as a presenting problem that is the result of multiple etiologies or comorbidities. In addition, the presenting problem may be the result of multiple environmental impacts on the subject and a lack of resources to cope with these forces. For example, cultural circumstances could play a central role; a misalignment between an individual's values and the values that are prominent in the broader culture could promote a distorted integration of the individual's internal self-definition and goals. Understanding the patient's self-narrative is thus an extremely important aspect of identifying and apprehending their presenting problem(s). iThe specialist's view also plays a significant role in the identification of dysfunction or illness; hence, their narrative and perspective on the patient's concerns and testimony are also crucial to clinical practice, which suggests that a methodology of atomism is far too simplistic to operate effectively in psychological and behavioral interventions.

c) The brain and behavioral health sciences neglect of the big picture: interaction

Humans and other animals perceive the world as structures that are uniform and organized. In the process of aging, humans become habituated to familiar objects, environments, and beings; some aspects of perceptual experience become less meaningful, and several reductions take place – for example, the perception of tridimensionality, color variation, and other complexities of visual experience that were once extremely salient become so familiar that we no longer take note of them but instead take for granted that these are the ways that visual stimuli appear to us. The process of acclimation promotes the formation of generalizations, fast learning, and efficient processing of sensory information, among other virtues.

The big picture problem intersects with the brain and behavioral health sciences' propensity to apply general biological categories to individuals in infinitely variable contexts. There are two good reasons for this: a) the fact that most humans follow a similar developmental trajectory, and b) the need to categorize specific constellations of signs and 'symptoms' as a disorder for every human, regardless of race, culture, individual idiosyncrasies, or the identity of the person who narrates the story of the dysfunction.

Nonetheless, it is paramount to remain aware of the danger of oversimplifications. We can start by recalling that in our generalizations, there are several levels of complexity involved, including different levels of abstraction, and the overall size of the scale we are taking – that of humanity itself – is undoubtedly enormous. Often, brain and behavioral health sciences seem to forget that any given cognitive dysfunction, psychiatric disorder, aberrant behavior, or similar phenomenon is inevitably immersed into a recursive system like this:



Therefore, most human beings undergo a similar process of development that includes many of the same propensities and limitations from person to person. It is also true that predictable patterns are likely to occur when a cognitive dysfunction, a psychiatric disorder, or an aberrant behavior develops. Nonetheless, to avoid falling into the "small picture" problem, meaningful cultural differences and individual subtleties should be retained in the form of unique narratives.

II. An out-of-date notion of 'personality'

Although there is no complete definition of personality, it can be understood as the compound of characteristics an individual manifests. An individual's personality is the singular and differentiated way of thinking, feeling, and behaving. Correspondingly, personality is influenced by experiences, environmental life situations, and inherited characteristics. According to the American Psychological Association (2020), the two main areas of the study of personality focus on understanding individual differences in particular personality characteristics and understanding how the various parts of a person come together as a whole.

Current brain and behavioral health sciences work with an outdated notion of personality brought to them by Psychiatry. Psychiatry considers personality a trait in the subject, a finished and completed product, and a stationary character. This can be clearly attested looking through the *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition* (2013) (DSM-5). The DSM-5 defines a personality disorder as "a way of thinking, feeling and behaving that deviates from the *expectations of the culture*, causes distress or problems functioning, and lasts over time" (American Psychiatric Association, my emphasis).[1] It declares that personality disorders are long-term patterns of behavior and *inner experiences* that differ significantly *from what is expected*. Personality disorders start in early adulthood and cause distress or problems in functioning. They can affect at least *two* of the following areas: way of thinking about oneself and others; way of responding emotionally; way of relating to other people, or; way of controlling one's behavior.

According to the DSM-5, there are 10 specific types of personality disorders: 1) Antisocial personality disorder (a pattern of disregarding or violating the rights of others, not conforming to social norms, repeatedly lying or deceiving others, or acting impulsively); 2) Avoidant personality disorder (exhibiting extreme shyness, feelings of inadequacy, and extreme sensitivity to criticism); 3) Borderline personality disorder (instability in personal relationships, intense emotions, poor self-image, and impulsivity. He/she may have repeated suicide attempts, display inappropriate, intense anger or have ongoing feelings of emptiness); 4) Dependent personality disorder (urgent need to be taken care of and submissive and clingy behavior, difficulty making daily decisions, or may feel uncomfortable or helpless when alone); 5) Histrionic personality disorder (excessive emotion and attention seeking, uncomfortable when they are not the center of attention, rapidly shifting or exaggerated emotions); 6) Narcissistic personality disorder (need for admiration, grandiose sense of self-importance, sense of entitlement, lack of empathy for others); 7) Obsessive-compulsive personality disorder (a pattern of preoccupation with orderliness, perfection and control, over focus on details or schedules, excessive work not allowing time for leisure); 8) Paranoid personality disorder (being suspicious of others and seeing them as mean or spiteful); 9) Schizoid personality disorder (being detached from social relationships and expressing little emotion, not caring about praise or criticism from others); 10) Schizotypal personality disorder (being very uncomfortable in close relationships, excessive social anxiety, distorted thinking and eccentric behavior) (American Psychiatric Association).[2]

The causes of personality disorders, according to the *American Psychological Association*, are genetics, parenting, and peer influences. They are grouped into five general categories:

1. Genetic factors, e.g., a gene mutation or abnormal interactions among genes.
2. Childhood traumatic experiences.
3. Verbal abuse, e.g., screaming at children or threatening to send them away.
4. High reactivity, e.g., the subject has high sensitivity to light, noise, texture, and other stimuli.
5. Peer interactions, e.g., peers can help either to prevent or bolster the development of person-

[1] American Psychiatric Association, 25 December 2020: https://www.psychiatry.org/patients-families/personality-disorders/what-are-personality-disorders

[2] American Psychiatric Association, 25 December 2020: https://www.psychiatry.org/patients-families/personality-disorders/what-are-personality-disorders

ality disorders (American Psychological Association).[3]

Several implications derived from this official canon are interesting. We will only mention three of them. First, when the DSM-5 refers to the way the subject "thinks about him/herself," "responds emotionally," or "controls" his/her behavior and how he/she "relates to other people," it is clear that all these amount to the way the subject narrates the world to him/herself. Second, when asserting that a personality disorder is a deviation "from the expectations of the culture," i.e., "what is expected," the definition provided by the DSM-5 is unintendedly opening the door to the ways the subject thinks, feels, and narrates his/her own world. Third, when the current DSM-5 states that "a person's personality typically stays the same over time" (DSM-5), it is emphasizing the fixed, finished/completed, and static aspects of personality.[4]

What is more, the DSM-5 states that diagnosis of a personality disorder "requires a mental health professional looking at long-term patterns of functioning and symptoms" (American Psychiatric Association).[5] It also states that people under 18 are typically not diagnosed with personality disorders because their personalities are still developing. These last points entail two relevant implications. First, based on the patient's symptoms, interviews, narratives, and procedural criteria, the mental health professional will build a story that puts all of this information into context, and therefore reach a diagnosis. Second, in people under 18, personality has not yet finished developing, which raises the question of when this point really is reached.

III. Challenging the brain and behavioral health sciences' notion of personality

a) Updating the notion of personality: Constructivist Approaches

In this section, some aspects of the notion of personality will be called into question, namely, the *finished product* conceptualization and its static character. We can start by recalling that sociocultural and constructivist accounts of personality and emotions are not novel. More than 30 years ago, insightful social construction descriptions of the subject were introduced which emphasized that subjects possess an assortment of traits, emotions, and conceptualizations that define themselves.

In opposition to the self-contained, static, and fixed accounts of personality, Armon-Jones (1985) instead asserted that emotions are responses to the environment "presented through specific modes of social organization, normative expectations, beliefs and values" (Armon-Jones 1985: 1). Additionally, she claimed that emotions are prescriptible and that "prescribed emotion rituals are not merely criteria for, but also ontologically constitutive of the agent's emotion feelings" (Armon-Jones 1985: 20).

An important follow-up question would then be, what about inappropriate emotions? In this respect, the author defended the claim that inappropriate emotions are not warranted by the evoking object or situation. For example, when exhibiting the wrong degree of emotional intensity – e.g., too little grief at a funeral – emotions are mediated by the particular beliefs, desires, attitudes, and idiosyncratic intentionality of the individual, but also by the social representation of the object or situation (Armon-Jones 1985: 16).

Under this radical constructivist perspective, emotions cannot be autonomous, but judgments are made in reference to some normative standard of social accuracy/inaccuracy in the emotional response. Similarly, deviant or inappropriate emotions can be explained as cultural conflicts of values (Armon-Jones 1985: 17–20).

As all of this serves to indicate, personality is not a *finished*, static, self-contained product but a social construction. This approach to personality includes the following features: a) cultural and narrative bias impacts the subject's identity formation and self-definition, and b) social normativity introduces judgments of the appropriateness of the individual's emotional responses.

b) Updating the notion of personality: cultural-environment and the self

Moving from Psychiatry's notion of personality to the constructivist and environmental approaches is not only desirable but also a priority. This shift implies, among other things, acknowledging the existence of:

1. A social normativity judging the appropriateness of the individual's emotional responses.
2. A cultural bias in the subject's identity formation and self-definition.

Offering an alternative to the view that the subject's self-identity is a finished product and highlighting the

[3] American Psychological Association, 25 December 2020: https://www.apa.org/topics/personality/disorders-causes
[4] DSM-5: www.dsm5.org
[5] American Psychiatric Association, 25 December 2020: https://www.psychiatry.org/patients-families/personality-disorders/what-are-personality-disorders

impact of narratives on the formation of the personality, we can quote Miller et al. (1990). They claimed the self is a social construction developed in early childhood through narratives and face-to-face interactions. They focused specifically on personal storytelling among diverse cultural groups intersecting the narrative self and face-to-face interactions. Characteristically, they noticed young children claim others' experiences as their own, i.e., they appropriate others' stories and narratives about themselves for the purposes of constituting their self-identity.

In a similar vein, Polkinghorne (1991) asserted that one's self-concept, or personal identity, is formed by adaptations of one's cultural stories and myths. Significantly, he noticed that under stressful conditions, a self-narrative might decompose, producing anxiety and depression. An appropriate treatment would thus be to assist in the reconstruction of a meaning-giving narrative of self-identity.

Nelson and Fivush (2004) have also presented a developmental perspective of the 'self.' They emphasize the centrality of autobiographical memory in the personal construction of an individual's own life experiences. The processes they call the emergence and development of the self are achieved through the contributions of multiple faculties and abilities, which include necessary memory abilities, language, narratives, adult memory talk, temporal awareness, and understanding of the self and others. They asserted that these components contribute to the timing, quantity, and quality of personal memories from the early years of life.

Overall, these approaches have a couple of essential assets. They allow us to account for the cultural differences and similarities among subjects. They also allow us to incorporate and highlight the role of narratives, which is paramount in contemporary psychiatry. Since, on these models, personality is not conceptualized as a completed product, further inquiries can be pursued regarding the nature and relationship of personality and person narratives as they change over time as the subject ages..

c) Updating the notion of personality: Enactivism

The relevance of narratives in the brain and behavioral health sciences is testified by enactivism in cognitive sciences. The ruling creed in psychiatry used to be that mental disorders stem exclusively from functional brain disorders, and therefore, mental disorders should be understood and treated as brain disorders (Insel and Quirion 2005: 2221). To reference Krueger (2019): observable behavior was not part of the ontology of the disorders.

In contrast, according to enactive views, moods and emotions are reducible to physical processes – perceptual, autonomic, somatovisceral, chemical and hormonal, motoric, etc. – spanning brain, body, and (sometimes) world (Colombetti 2017; Colombetti & Roberts 2015). According to one of the founding fathers of enactivism, the way a subject thinks, feels, and perceives consists in an "enactment of a world and a mind on the basis of a history of the variety of actions that a being in the world performs" (Varela et al. 1991: 9).

Some of the best examples to illustrate the pertinence of enactivism are Autism Spectrum Disorders (ASD). Subjects with autism are often described as subjects who exhibit a series of cognitive and physical behaviors such as: a failure to integrate visual holistic qualities of objects in the way that neurotypical individuals mostly do; abnormalities in eye contact; difficulties in the triad of social-communicative-imaginative skills; altered interpersonal and interactive capacities; exaggerated gestures; stereotyped behaviors; repeated rocking or tics and proprioceptive processing difficulties, among several other characteristics that have been documented in the last decades (Atkinson 2009; Blanche et al. 2012; Doan et al. 2013; Eigsti 2013; Frith 2003; Leary and Donnellan 2012; Mottron et al. 2006). When we consider the set of ASD characteristics through the lens of enactivism, we formulate a coherent narrative that leads to both a diagnostic classification and relevant strategies for intervention and therapy for ASD patients.

As with ASD and most developmental disorders and cognitive dysfunctions, it is reasonable to assume that cognitive activities are irreducibly enacted and even embodied. In other words, they are made up of processes looping through brain, body, and world, as a result of our ecological dealings with people and things around us (Krueger 2019: 2). Altogether, these considerations have important implications for clinical approaches to disorders in the brain and behavioral health sciences, and more concretely, for individual therapeutic encounters.

Arriving at a diagnosis is possible through consideration of both concrete, empirical, measurable criteria – e.g., the patient displays compromised brain activity in specific sensorimotor circuits – and the narrative of an expert (a therapist, physician, psychiatrist, or psychologist) who describes the patient's specific qualitative characteristics – e.g., diminished facial expressions, a lack of experiential interaction, etc.. Thus, in addition to analyzing the patient, the expert is experiencing the patient. To quote Stern:

"We naturally experience people in terms of their vitality. We intuitively evaluate their emotions, states of mind, what they are thinking and what they really mean" (Stern 2010: 3–4).

The embodied and expressive features that can be observed, plus the clinical encounter's qualitative character, are part of the diagnostic and therapeutic process (Krueger 2019: 19). The behaviors we see in others have a kinematic signature, are environmentally situated, are regulated normatively by sociocultural practices, and exhibit specific details of concrete situations, occurrence patterns, etc. All those aspects are the ones we must incorporate in our narratives about others in all the brain and behavioral health sciences.

Conclusion: The fitting and circumscribed role of narratives

This paper has examined the role of narratives in the contemporary brain and behavioral health sciences, focusing in particular on the implications for clinical practice. The first goal was to explain why narratives have been underestimated in the literature. Second, we have underscored the need to update the notion of personality in psychiatry, related fields, and brain and behavioral health sciences in general.

A final remark concerning the limits of narratives is due. Before concluding, a pertinent question remains: Are we willing to advocate the priority of narratives in the brain and behavioral health sciences, and to what extent?

This work is not a plea for narratives at the cost of current ongoing practices in psychiatry and the behavioral health sciences. A narrative/story/description itself cannot serve as a substitute for evidence-based practices in the brain and behavioral health sciences. However, narrative has a crucial role to play. This role is not jeopardized by the existence of multiple narratives or the possible question of, 'who narrates the story better?' This becomes clear when we understand that the function of narrative is to integrate the many features of a complex phenomenon, which include concrete evidence and several (potentially contrasting) stories: the specialist's observations, the patient's self-description at different stages of life, the cultural attitudes toward the disorder/dysfunction, the scientific community's typology, etc.

Narratives play an active role in the diagnosis and treatment of many conditions. Often, the subjective elements of narrative, such as a clinician's experience of the patient or the patient's sensations and inner states, can indicate something related to diagnosis or treatment is amiss. Ignoring the epistemic value of narrative may thus pose very real consequences for patients and their families. Incorporating narratives into clinical and theoretical systematization is the clear solution to this. Therefore, some minimum criteria must be established to distinguish narratives that ought to be considered epistemically-valuable and authoritative. These criteria are as follows:

1) Consistency. Narratives should be sufficiently consistent with diagnostic practices in the brain and behavioral health sciences, i.e., with DSM psychiatric criteria, neuroimaging results, etiology signs, clinical symptoms, and the like. Narratives could exhibit significant challenges to one or some standards, forcing us to review a case. Still, when the patient or the expert description contradicts other evidence, the narrative itself is what we should reconsider.

2) Explanatory power. An acceptable narrative should minimally connect a state or event A with a state or event B or C in a consecutive fashion. This implies that the report should sequentially track episodes. A narrative that discards events randomly should be weighed with caution.

3) Coherence and constancy. A narrative must track events that occur in reality and in other related narratives. The narrator's account should be continuous with the patient's and with the expert's description.

A narrative that meets these conditions – that is consistent, explanatory, coherent, and constant – can be taken as consequential, indicative, robust, and appropriate in the context of clinical practice.

Future directions regarding the proper uses of narratives in the brain and behavioral health sciences will include documenting:

a) The cognitive-affective bundle in the narrator's performance, i.e., the cognitive and affective influences at play in the patient's (Challita 2016), the specialist's, and the scientific community's narratives.

b) Some brain development facts, e.g. how the human brain is half-developed at birth, taking twelve to fourteen years to fully develop; the continuous maturation process of the frontal lobe until the third decade of life; etc. (Striedter 2005; Striedter & Northcutt 2020). Such facts are necessary constraints to implement in any diagnostic consideration, since they are likely to influence narrative perspectives (in the self, in the specialist, and in the experts' community).

c) Crossed comparative analyses of differential modifications in the patients' narratives due to dramatic changes in the cultural environment.

d) The emergence of human autobiographical memory in the past (Nelson & Fuvish 2004) and its development across the preschool years (from basic memory abilities, language, temporal understanding, and understanding of self and others) up to adult memory talk, the subsequent changes of personality, and the concomitant modification in narratives as the subject ages.

e) A historical revision, by category, for each of the ten current DSM-5 personality disorders. For instance, the case of social phobia, "how shyness became an illness" (Lane 2006), could fit into more than one DSM-5 personality disorder type. Ideally, it could be the hallmark of (2) Avoidant personality disorder (exhibiting extreme shyness, feelings of inadequacy, and extreme sensitivity to criticism). But extreme shyness could also characterize other disorders. It could be present in: (8) Paranoid personality disorder (being suspicious of others and seeing them as mean or spiteful), (9) Schizoid personality disorder (being detached from social relationships and expressing little emotion, not caring about praise or criticism from others), and (10) Schizotypal personality disorder (being very uncomfortable in close relationships, excessive social anxiety, distorted thinking, and eccentric behavior). There are such broad possibilities for classification and exclusion criteria that are quite unclear in the DSM-5 that the problem of diagnosis and classification rather seeds doubts.[6]

f) The contrasts between the co-constructive and the sociocultural accounts.

g) Enactive views and their connection with narratives, but also the rest of the 4E approaches, i.e., the Embedded, Embodied, and Extended ones.

h) Expressive features and bodily manifestations of psychopathologies. This is a priority when such manifestations exhibit specific kinematics in addition to other possible hidden psychic dimensions to support diagnostic and therapeutic interventions (see Krueger 2019 and her discussion on Forms of Vitality).

i) The defense of the Direct Social Perception. This may be useful by helping us to know more about the phenomenal character of disorders and how we can perceive them in others.

[6] This point reminds us of Karl Popper's philosophy of science's old attack on psychoanalysis. Popper's (1935) criterion to demarcate science from pseudoscience was falsifiability. He claimed genuine empirical sciences specify the conditions under which they would readily accept their falsifiability. Pseudoscientific sciences, on the contrary, cannot be falsified.

References

Allen C, Neal J (Spring 2020). Teleological Notions in Biology. In *The Stanford Encyclopedia of Philosophy,* ed. Edward N. Zalta. https://plato.stanford.edu/archives/spr2020/entries/teleology-biology/.

American Psychological Association (25 Dec 2020), https://www.apa.org/topics/personality.

Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition-DSM 5 (2013). American Psychiatric Association. Washington: American Psychiatric Publishing (25 Dec 2020). www.dsm5.org

Andersen H (2020) Reductionism in the Biomedical Sciences. In *The Routledge Companion to Philosophy of Medicine, ed.* M. Solomon, J. R. Simon, H. Kincaid, 81–89. New York and London: Routledge, Taylor and Francis.

Armon-Jones C (1985) Prescription, Explication, and the Social Construction of Emotion. *Journal for the theory of social behavior*, 15(1): 1–22. Doi:10.1111/j.1468-5914.1985.tb00042.x

Ariew A (2002) Platonic and Aristotelian Roots of Teleological Arguments. In *Functions*: *New Essays in the Philosophy of Psychology and Biology,* ed. A. Ariew, R. Cummins, M. Perlman, 7–32. New York: Oxford University Press, Clarendon Press.

Atkinson A P (2009) Impaired recognition of emotions from body movements is associated with elevated motion coherence thresholds in autism spectrum disorders. *Neuropsychologia,* 47(13): 3023–3029.

Averill J R (1985) The social construction of emotion: With special reference to love. In *The social construction of the person,* ed. K. J. Gergen, K. E. Davis, 89–109. New York: Springer-Verlag.

Blanche E I, Reinoso G, Chang M C, Bodison S (2012) Proprioceptive processing difficulties among children with autism spectrum disorders and developmental disabilities. *The American Journal of Occupational Therapy: Official Publication of the American Occupational Therapy Association,* 66(5): 621–624.

Carnap R (1934) *The Unity of Science*. London: Kegan Paul, Trench, Trubner & Co.

Challita M (2016) From empathic mind to moral behaviour: the 'who', 'why' and 'how'. *Medicine, Health Care and Philosophy,* 19: 517–522. Doi: 10.1007/s11019-016-9694-5.

Colombetti G (2014) *The Feeling Body: Affective Science Meets the Enactive Mind*. Cambridge, MA: MIT Press.

Colombetti G (2017) Enactive Affectivity, Extended. *Topoi*, 36(3): 445–455.

Colombetti G, Roberts T (2015) Extending the extended mind: The case for extended affectivity. *Philosophical Studies*, 172(5): 1243–1263.

Cornish-Bowden A (2006) Putting the Systems Back Into Systems Biology. *Perspectives in Biology and Medicine*, Autumn, 49(4): 475–89.

Doan M, Fenton A (2013) Embodying Autistic Cognition: Towards Reconceiving Certain 'Autism-Related' Behavioral Atypicalities as Functional. In *The Philosophy of Autism, ed.* J. L. Anderson, S. Cushing, 47–71. New York: Rowman and Littlefield.

Eigsti I M (2013) A Review of Embodiment in Autism Spectrum Disorders. *Frontiers in Psychology*, 4(224): 1–10.

Frith U (2003). *Autism: Explaining the enigma*. Hoboken: Wiley-Blackwell.

Hernández-Chávez P (2019) Blinded by Biology: Bio-inspired Tech-Ontologies in Cognitive Brain Sciences. In: *Bio-inspired Information and Communication Technologies*, ed. A. Compagnoni, W. Casey, Y. Cai, B. Mishra, vol 289, 58–69. Springer. DOI: https://doi.org/10.1007/978-3-030-24202-2_5.

Hernández-Chavez P, García-Campos J (2020) Six Classic Tools for Conceptualizing Cognitive Dysfunctions: Why 'Classic' means 'Timeless'. Under review.

Huneman P, Wolfe C (2010) The Concept of Organism: Historical, Philosophical, Scientific Perspectives. *History and Philosophy of the Life Sciences* 32(2–3): 147–154.

Insel T R, Quirion R (2005) Psychiatry as a clinical neuroscience discipline. *JAMA, the Journal of the American Medical Association*, 294(17), 2221–2224.

Krueger J (2019) Enactivism, other minds, and mental disorders. *Synthese:* 1–25.
https://doi.org/10.1007/s11229-019-02133-9.

Laubichler M D (2000) The Organism is Dead. Long Live the Organism! *Perspectives on Science*, 8(3), 286–315.

Lane Christopher (2006) How Shyness became an Illness. A Brief History of a Social Phobia. *Common Knowledge* 12: 3. DOI 10.1215/0961754X-2006-004.

Leary M R, Donnellan A M (2012) *Autism: Sensory-movement Differences and Diversity*. Cambridge: Cambridge Book Review Press.

Miller P J, Potts R, Fung H, Hoogstra L, Mintz J (1990) Narrative practices and the social construction of self in childhood. *American Ethologist*, 2(17): 292-311.doi.org/10.1525/ae.1990.17.2.02a00060.

Mottron L, Dawson M, Soulières I, Hubert B, Burack J (2006) Enhanced perceptual functioning in autism: An update, and eight principles of autistic perception. *Journal of Autism and Developmental Disorders*, 36(1): 27–43.

Nelson K, Fivush R (2004) The Emergence of Autobiographical Memory: A Social Cultural Developmental Theory. *Psychological Review*, 111(2): 486–511 .doi:10.1037/0033-295x.111.2.486.

Nicholson D J (2014) The Return of the Organism as a Fundamental Explanatory Concept in Biology. *Philosophy Compass*, 9: 347–359.
doi: 10.1111/phc3.12128

Perlman M (2009) Changing the Mission of Theories of Teleology: DOs and DON'Ts for Thinking about Function. In *Functions in Biological and Artificial Worlds: Comparative Philosophical Perspectives, ed.* Ulrich Krohs and Peter Kroes. Vienna Series in Theoretical Biology.Cambridge, Massachusetts: The MIT Press.

Polkinghorne D E (1991) Narrative and self-concept. *Journal of Narrative and Life History*, 2–3: 135–153. doi.org/10.1075/jnlh.1.2-3.04na

Popper K (1935/1959). *Logic der Forschung zur Erkenntnis*. Vienna: Springer. English translation: *The Logic of Scientific Discovery*. Hutchinson, London.

Schaffner K F (2006) Reduction: The Cheshire cat problem and a return to roots. *Synthese*, 151(3): 377–402.

Schaffner K F (2011) Reduction in biology and medicine. *Philosophy of Medicine*, 16: 137.

Stern D (2010) *Forms of Vitality: Exploring Dynamic Experience in Psychology, the Arts, Psychotherapy, and Development*. Oxford: Oxford University Press.

Striedter G F (2005) *Principles of Brain Evolution.* Massachusetts: Sinauer Associates.

Striedter G F, Northcutt R G (2020) *Brains Through Time: A Natural History of Vertebrates*. Oxford University Press.

Varela F J, Thompson E, Rosch E (1991) *The Embodied Mind: Cognitive Science and Human Experience*. Massachusetts: MIT Press.

Webster G, Goodwin B C (1982) The Origin of Species: A Structuralist Approach. *Journal of Social and Biological Structures*, 5(1): 15–47.
DOI: 10.1016/S0140-1750 (82)91390-2.

NARRATIVES IN HEALTH CARE: A CASE FOR PSYCHOEDUCATION DRAWING ON THE BIOPSYCHOSOCIAL MODEL

Isabella Sarto-Jackson
Konrad Lorenz Institute for Evolution and Cognition Research
Klosterneuburg, Austria
e-mail: isabella.sarto-jackson@kli.ac.at

Abstract:
The role of narratives in clinical practice has long been underappreciated. This disregard is largely due to an overemphasis on reductionist interpretations of disease causes based on the primacy of the medical model of disease. This way of thinking has led to decontextualizing symptoms of disorders from patients' lives. More recently, however, healthcare professionals have turned towards a biopsychosocial model that reintroduces sociocultural and psychosocial aspects into clinical diagnosis and treatment. To this end, narrative approaches have been increasingly explored as alternative diagnostic and therapeutic tools. Central to the narrative approach is the avoidance of pathologizing language that usually focuses on deficiencies. Instead, patients' narratives are co-constructed and co-created together with the clinician or therapist to transform them into empowering stories about healing. To make narratives accessible and transformable for the patient, psychoeducational methods can be used to translate scientific and medical knowledge about the disease into stories described in everyday language that resonate with the patient's own life stories. Consequently, psychoeducational narratives enhance the patient's competence in coping with a physical or mental illness and re-contextualizing symptoms, and prompt an increased compliance with therapies.

Key words: narratives, medical model of disease, biopsychosocial model, psychoeducation, narrative approach, narrative therapy.

From the medical model of disease to an updated biopsychosocial model of health

Over the last decades, the medical model of disease (Laing 1971) has become the most widely accepted model in modern scientific medicine, defining basic assumptions about standardized assessments of illnesses, the underlying causes of illnesses, and their remediation. In practical terms, the medical model of disease necessitates a set of procedures that submit to objectivist and causal-mechanistic principles (Sarto-Jackson 2018) that are largely based on the biostatistical theory (Boorse 1975, 1977). The biostatistical theory argues that disorders are no more than an organ's or a physiological process' biological dysfunction that can be clearly identified by a measurable statistical deviation from its typical functioning. Most healthcare workers, including mental healthcare workers, have come to embrace the medical model of disease (Ludwig 1975, p. 603). Accordingly, psychiatric diseases and mental illness have been defined as biological brain dysfunctions, either biochemical or neurophysiological in nature, making all psychopathologies basically physiopathologies. The biostatistical view that informs the medical model has marginalized alternative approaches such as social models, holistic models of the alternative health movement, and the recovery models of mental health.

Despite its success for the diagnosis and treatment of many somatic diseases, increasingly more healthcare workers and patients have run into limitations when fully subscribing to the medical model, especially when dealing with mental illnesses. Today, most human diseases are considered complex polygenic and multifactorial disorders that are caused by a multitude of genetic and environmental influences acting in concert. Since this seems to hold true for almost all somatic diseases, it will in all likelihood also hold true for psychopathologies. If diseases are influenced by life-history trajectories, environmental factors, and epistasis, all causing significant interindividual variabilities between apparently "equivalent" diseases, it will be difficult to uphold the current view on the medical model of disease. The latter implies that diseases can be categorized by distinct and specific biochemical signatures which are by and large generalizable for all patients suffering from the "same" disease. In fact, so far there has been little evidence for reliable biomarkers for any mental disorder (Deacon 2013). Since the medical model provides the ultimate rationale for therapeutic measures, in particular the choice of psychiatric drugs, the supremacy of pharmacotherapeutic interventions for any mental disorder is challenged when additional factors like life-history trajectories, environmental factors, and epistasis are taken into account.

Central to the medical model is the notion of typical functioning and the assumption of "normal" physiological and neurophysiological processes that are defined to be characteristic of a healthy person. Accordingly, deviations from normal functioning are believed to be expressed as symptoms that can be treated by psychiatric drugs. Drugs then reverse, or partially reverse, the neurophysiological "abnormality" that gives rise to the symptoms. This model of thought has been termed the disease-centered model (Moncrieff & Cohen 2005). Interestingly, the disease-centered model of drug action has also come under criticism recently. Rather than psychoactive drugs correcting an abnormal physiological state in a targeted manner, it may well be that drugs exert their effects in a systemic manner by shifting the organism into some altered mental states. Such altered states may indeed (partially) suppress symptoms, yet not by targeting the underlying causes of the mental disorders, but by incidentally mobilizing systemic feedback loops of the organism. The latter, alternative concept has been termed the drug-centered model (Moncrieff 2008). This novel conceptual model predicts that psychiatric drugs globally induce altered mental states that not only mitigate disease symptoms, but concomitantly modify a patient's overall thinking, feeling, and behavior. These predictions are in good agreement with clinical observations of drug effects. However, under the traditional conception of the disease-centered models, such alterations in thinking, feeling, and behavior were considered side effects of the active pharmaceutical ingredient.

Due to the challenges that psychopathological research and psychopharmacotherapy is facing, health experts have increasing looked for alternative therapeutic measures – not least because the medical model of diseases has fostered the separation of a holistic definition of illness and an objectivistic (predominantly biostatistical) definition of disease. This separation has contributed to a startling disjunction between the patient as a person and the patient's disorder. As a consequence, the patient's role within the medical model has been increasingly perceived as that of a passive sufferer rather than that of an affected person with a distinct scope of action. The disease-centered model of psychopharmacotherapy and its underlying reductionist hypothesis about how drugs exert their effects as well as the disregard of holistic aspects in mental healthcare have led to a strong critique of the medical model with respect to psycho-pharmacotherapeutic interventions. The concept of the medical model is heavily burdened with a reductionist view in which measurable biological parameters have become the ultimate criteria defining disease, while at the same time the patients' life histories and sociocultural effects are usually ignored. As a consequence, physicians and psychiatrists are often facing the paradox that some people with positive laboratory findings are considered in need of treatment when in fact they are feeling well, while others feeling sick are dismissed as having no "disease" and must therefore feel well. Grappling with these controversies and having become more and more aware of the limitations of the medical model of disease, many healthcare workers now embrace a biopsychosocial model, an alternative model of health that was developed in the late 1970s (Engel 1977). The latter was significantly influenced by general systems theory providing a shift from linear cause and effect to circular causality (Bertalanffy 1972).

In general systems theory, all levels of organization are interconnected in a hierarchical relationship so that a change at one level affects changes at the other levels. Some fundamental principles operate commonly at all levels of organization, while others are unique for each level. For the biopsychosocial model this would refer to different levels of organization, including biomolecules, cells, organs, the organism, the person, the family, the society, and the biosphere. Adopting systems theory in medicine could mitigate the holist–reductionist dichotomy and foster exchange and collaborations across different fields of healthcare, research, and social work. Based on the biopsychosocial model, health cannot simply be defined by an organism's statistically normal functioning, but is considered as the organism's competence to vanquish pathogenic factors on many levels in an autoregulatory manner (e.g., to cope with bacteria, viruses, toxins, adverse psychological, sociocultural, or economic conditions). Following this line of argument, a disease will occur if an organism's autoregulatory competence is insufficient or systemic – that is, if its organismal and adherent (e.g., socio-environmental) feedback loops are overburdened. In other words, the existing medical model of disease does not suffice because it limits itself to a few isolated levels of biological organization. In contrast, the biopsychosocial model also takes into account the patient's developmental trajectory, the social context of living, the societal system that scaffolds how to deal with the adverse effects of illness, and the physician's role and that of the healthcare system. Thus, the biopsychosocial model aims at putting back sociocultural, ethnomedical (Fabrega 1975), and psychosocial perspectives into an overall framework. This naturalistic framework draws from the patient's life history, ontogenetic factors, and social niche, thereby giving more genuine room for social, psychological, and behavioral dimensions. At the same time, the biopsychosocial model provides a differentiated approach for understanding the determinants of disease and consequently arrives at other, additional treatments and modes of healthcare.

While the biomedical model has encouraged bypassing the patient's verbal account by placing greater reliance on technical procedures and laboratory data, the biopsychosocial model acknowledges the complexity of a disease, of the diagnosis, and of the resulting therapeutic interventions. Thus, the biopsychosocial model of health expects the physician's or psychiatrist's professional knowledge and skills to span the social, psychological, and biological realm for diagnosing the patient's condition and to be able to resort to either of the three realms (or use a combination of those) to decide and act upon when treating the patient. Along this line of thinking, medical examinations or psychiatric assessments require the correlations of laboratory measurements and clinical data based on biomedical methods as well as on clinical data collection (e.g., high-level interviewing skills), on one hand. On the other hand, they also require an understanding of the psychological, social, and cultural dimensions of how patients communicate symptoms of disease (Engel 1975) and how they perceive their own role as a patient.

Narratives as tools in clinical practice

In philosophical literature, work on the theory of narrative has largely concentrated on literary (especially fictional) and – more recently – historical narratives. Yet, narratives broadly construed are means by which a story is told, whether fictional or not and regardless of the medium. A particular focus lies on temporal and causal relations of crucial events. Narratives have been elevated because of the value attributed to a person's life and acts when those acts conform to a conception of that life as exemplifying narrative. Of course, not all action descriptions would be called narratives. For example, traditionally, psychological or sociological descriptions of behavior and social interactions of individuals or groups are usually not understood as narratives. Also, accounts of scientific, mathematical, or physicochemical principles are not considered narratives. However, boundaries between narration narrowly construed and other means of inter-agency communication can be blurred. In general, narratives presuppose a narrator mediating between "audience" and action (i.e., behavior), thus rendering a distinct contrast to interpersonal discourse where action is directly performed. In addition, narratives are intended to have certain effects on the audience – thus, in clinical settings the "audience" intentionally includes the narrator to benefit from the approach described below.

The role of the patient-physician dialogue has long been underappreciated, with many largely ignoring the worth of narratives in clinical practice as valuable tools for conceptualization, diagnostics, and treatment of disorders. This disregard can be attributed to a decontextualization of patients' lives by overemphasizing biological and reductionist interpretations of disease symptoms based on the primacy of the medical model of disease. More recently, by turning towards the biopsychosocial model of health, psychiatrists have integrated alternative tools that also capture the psychological and social aspects of the patients' lives. In particular, narratives have become an important tool in psychiatry. Healthcare professionals now often enrich traditional medical-oriented practices – interviewing, diagnosing, and treatment – with narrative practices. By doing this, physicians draw on the insights from evidence-based medicine, genetics, molecular biology, neuroimaging, and psychopharmacology, while acknowledging that these fields of science are not outside of the patient's psychosocial context and can therefore be complemented by narratives. Hence, physicians who also call themselves "narrative medical clinicians stand at the intersection of the world of stories and the world of scientific categorization" (Carrey 2007, p. 78).

The aim of using narrative tools in psychiatry is to collaboratively work with patients, utilizing their narratives in clinical practice, research, and education as a way to foster the patients' ability to cope with their physical or mental illnesses. By focusing on individual stories of patients, the relational and psychological dimensions that co-occur with the disease are addressed and transformed into co-constructing or co-creating empowering stories about healing. This medical approach enriches the collaborative patient-doctor work by validating the patients' experiences and recognizing the patients' knowledge about their own condition (White & Epston 1990). As a consequence, a narrative approach can avoid conflicts about patients' compliance. Such conflicts can emerge when therapeutic interventions solely rely on the traditional expert status of the clinician, while patients challenge this view by considering themselves as "experts" about their "reality." Central to the narrative approach is the co-construction of a salutary story by both doctor and patient (Parry & Doan 1994), while acknowledging that truth and knowledge are not absolute, prescriptive, and paradigmatic, but biographical and interpretive (Launer 1999). At the same time, narratives can be "owned" by the patient and thus invoke the patient's committed involvement by reducing uncritical acceptance of the psychiatrist's expert knowledge. A particular challenge is the fact that this practice requires the narrative clinician to be versed in biomedical knowledge of mental or behavioral disorders as well as in the art of translating this scientific and medical knowledge into a story that resonates with clients' own stories, which necessitates creativity and self-reflection in the physician.

Narratives in the form of self-reports by patients can help medical professionals to integrate symptoms into a wider conceptual framework. In fact, using narratives in clinical practice may contribute to a deepened understanding of the patient's condition that goes beyond the prevalent evidence-based practice and exceeds the mere observation of deviant behavior (Gurova 2013). But even more so, by utilizing the privilege that society and culture have invested in psychiatrists, the latter can co-author and transform the patient's narratives to promote healing. The healing process can be set in motion by exploring the social and life-historical origins of the problem, providing meaning, context, and perspective for the patient's predicaments, and unpacking conventional biomedical explanations. Consequently, narratives can instill dynamics and momentum to definite diagnoses of mental illnesses through self-reflection and can afford the formation of novel coping strategies in patients (Kirmayer & Crafa 2014).

Telling a different story

Narrative psychiatrists have questioned claims of objectivity in conventional psychiatric assessments in which biomedical discourses are given a privileged role over patients' stories (Simblett 1997). This critique is based on the observation that each stage of the assessment process (including the medical statement concerning the referral, the presentation of the medical problem, the patient's history, the extended family-personal history, the examination of the patient's mental status) is usually conducted to inadvertently focus on the patient's problems and deficiencies and is couched in the psychiatrist's language. When a biomedical language of "disorder" and "psychopathology" is used that draws on the medical model of disease and aligns with the categories outlined in the Diagnostic and Statistical Manual (DSM), it may obscure psychosocial or developmental dimensions. Moreover, some behaviors that fall under DSM classification can vary in "adaptability" (depending on gender, age, social status, etc.). Manifestations of such disorders may thus change depending on how the environment responds to them. If patients' stories are dissociated from their environmental, familial, or social context, looking at the medical diagnosis in isolation, clinicians run the risk of amplifying the problem they are trying to solve. Hence, classifying patients by applying psychiatric labels may contribute to the patients' marginalization and isolation as other people (including spouses, parents, friends) may strongly "believe" in biomedical labels, thereby causing diagnoses to become self-fulfilling prophecies (Launer

1999). This phenomenon was elaborately described by Ian Hacking arguing that mental disorders are so-called human kinds – in contrast to natural kinds – that undergo looping effects typical for scientific classifications (Hacking 1995). When diagnosed with a mental disorder, such medical labels can bring on changes in the self-concept and the behavior of individuals so classified. These changes can, in turn, lead to revisions in the initial descriptions of the mental disorder.

Thus, assessment and diagnosing represent challenges, but also opportunities. In contrast to physicians who subscribe to the medical model of disease, narrative psychiatrists avoid making explicit diagnosis statements, which may give normative labels to conditions that concomitantly bear non-medical judgments, including "normal" or "abnormal" behavior (Carrey 2007). Narrative therapists co-construct alternative, more hopeful possibilities with their patients by using metaphors that are more helpful than those of organic diseases. Rather than submitting to disease metaphors that reinforce helplessness and inevitability, narrative approaches are more transparent about therapy processes and the roles that patients as well as their social environment play. This procedure acknowledges the fluidity of disease categories, puts emphasis on multiple (biological and other) factors that co-create such conditions, and utilizes diagnoses in a tentative manner. In this way, reductionist models that decontextualize disorders from the patients' rich and complex lives are challenged, leading to different, often liberating opportunities for patients.

In practical terms, narrative clinicians invite patients to re-author their lives by drafting a preferred storied life. This act of relabeling serves to re-story the patient's identity in agreement with an identity story they prefer. This procedure also gives back agency to the patients rather than leaving them to grapple with an abstract disease metaphor that inevitably escapes their influence. The positive outcomes based on narrative approaches are supported by decade-long studies on the effects of psychotherapy understanding language as a social phenomenon as well as a profoundly human need (Gomez 1996). Language not only shapes us, how we act and interact as humans, but language also creates stories that can make us the person we intend to become. In this sense, narratives hitch on to our self, give our lives meaning, continuity, and coherence. Some may even go further and say that our very personal identities rest on narratives we form about ourselves. In this respect, the self is more than a mere psychologistic bundle of properties particular to a certain person, but is crucially tied to the agency that drives these properties. And as agents who construct the external self, we rely on narrating a story of our lives. In agreement with a theory of narratives, narratives must be more than just a sequence of events. Narratives require a causal chain that knits events together. Such a causal structure is usually provided by an agent. Assuming the patients to be the agents of their narratives in narrative approaches, they can be considered free in how to fit relevant life events into a causal structure of their own choosing. In doing so, a narrative approach gives them an inalienable autonomy, the right to narrate their own lives. This concept resonates with the self as a center of narrative gravity (Dennett 1992) that describes the self as the creation of a certain personal narrative, a sort of living novel. According to Dennett, we all, at times, confabulate, tell and retell ourselves the story of our own lives, being "inveterate and inventive autobiographical novelists" (Dennett 1992, p. 111). Support for this concept comes from neurobiological and neurosurgical findings of split brain patients who are missing parts of the full narrative which is not accessible to the nonlinguistic part of the brain (Gazzaniga 1992).

There is certainly more to the self – which thinks and feels – than just those inner processes that constitute an elaborate narrative that contributes to one's identity. The self requires an "inner point of view" that – usually at the level of subconsciousness – integrates different sensory, perceptual, and cognitive modalities and is interconnected with memory and awareness. In addition, the self relies on reflexivity (i.e., a mind that loops back), which makes it possible that the self is introspectively accessible and can provide a subjective perception of the ownership of thoughts. These aspects of self would require a much more elaborate analysis given the huge body of philosophical literature that spans several centuries. However, the self is not the matter of main concern here. Instead, the point I want to make is that humans use narratives to contour their self by telling their own history, analyzing their status quo, and narrating their own future. As far as we know, humans are the only creatures that can position themselves on such a temporal trajectory and thus adopt a metacognitive standpoint about their selves in time. If we can agree on this point, it should not be too far-fetched to conclude that changing narratives can strongly influence a person's perception of identity as part of its self, which makes the narrative approach in psychiatry an incredibly powerful tool.

From narratives to psychoeducation

It is well documented that the behavior of the physician and the relationship between patient and healthcare professional strongly determine the therapeutic success (Chipidza, Wallwork, & Stern 2015). Psychological factors can influence the patient's experience of the disease and modify the patient's reaction through psychophysiological processes (Benedetti 2013). Thus, the successful application of therapies is surprisingly tightly-linked to the physician's ability to influence and modify the patient's behavior in compliance with her health needs. In a nutshell, "the physician's role is, and always has been, very much that of *educator and psychotherapist"* (Engel 1977, p. 132; my emphasis).

First of all, narrative approaches aim at avoiding the use of "pathologizing language" in medicine and psychiatry that is focused on the patient's deficiencies. Such pathologizing language reinforces power imbalance in the doctor-patient relationship and reifies traditional positions in the power hierarchy in which the clinician "knows," while patients do not. When operating from a dominant psychiatric discourse based on the medical model of disease that usually correlates disorders or psychiatric or behavioral problems with "chemical imbalances" and "hardwired" genetic influences, patients may not only feel helpless but also ashamed or guilty. The latter is especially true if medication or traditional therapies do not bring the desired effects. Such feelings of shame or guilt may come from a misconception regarding the precision of drug action, as patients may assume drugs to be highly specific, on target, and uniformly effective (Moncrieff & Cohen 2005). Any failures of the applied interventions may then be interpreted by the patient as a confirmation of being a hopeless case due to his or her individual shortcomings. This conclusion may be particularly harmful if concomitantly other factors (social stress, emotional issues, other professional or intense systemic pressures, etc.) are overlooked. By avoiding a pathologizing language of deficiencies in narrative approaches, feelings of shame and guilt in the patient which often occur in traditional settings, can be minimized.

Narrative approaches concern all stages of psychiatric or therapeutic doctor-patient interactions – interviewing/assessing, diagnosing, and treating. In all stages of narrative therapy, the traditional, standard therapeutic format in which the dominant psychiatric discourse presupposes a kind of "truth" about people and their lives is overcome. Instead, a narrative approach prompts therapists to acknowledge their position of power and use this knowledge to create a therapeutic environment in which their own voices are no longer automatically privileged over their patients', but are offered to help the latter to equally trust their own voice.

One of the biggest challenges of narrative therapy is the requirement that the narrative clinician be versed in clinical and biomedical questions about brain processes and putative dysfunctions as well as in the art of communicating with patients in ways that enable therapeutic effects through narratives. While narrative approaches, as discussed here, usually lie in the professional domain of psychiatrists and physicians, narrative therapy is a form of psychotherapy based on therapeutic conversations, strictly speaking. This form of therapy seeks to enable patients to identify their values and mobilize skills associated with them. It provides the patient with knowledge about their abilities to live these values in order to effectively confront current and future problems. In this sense, this approach clearly challenges dominant discourses based on the medical model of disease. Narrative therapy is often used by psychotherapists who specialize in family therapy, and also by social workers who specialize in community work or schools and higher education (Nylund & Tilsen 2006; Winslade & Monk 2000).

As narrative therapy draws on the biopsychosocial model, it highlights developmental aspects that allow for a re-evaluation of hereditary as well as psychosocial factors within an individual's development. Thus, the theoretical basis lies in the interplay of the nature–nurture influences. By focusing on the developmental trajectories that are shaped by multiple biomedical, social, environmental, and economic factors, the narrative about a medical or case history becomes much more nuanced, more accessible to the patient, and open to being retold. Yet, the more severe the patient's suffering, the more difficult it will be to retell a story. The psychotherapist seeks to help the patient to co-author a new narrative about themselves by investigating the history of the client's values that are crucial for mobilizing skills and abilities for the purpose of the healing process. For a specific case of illness or disorder, the psychotherapist or clinician must translate scientific and medical knowledge into a story that resonates with the patient's own life stories.

To increase the patient's competence about how to handle his or her disease, it is necessary to have a basic understanding of the illness in general and of putative causes underlying his or her health problems. To this end, clinicians and psychotherapists have increasingly adopted psychoeducational methods to inform their patients. Psychoeducation is a systematic and structured transfer of scientifically sound knowledge about somatic or mental illnesses. It has been used in evidence-based therapeutic interventions for psychopathologies (e.g., schizophrenia) and somatic illnesses (e.g., cancer; chronic pain). In psychoeducational interventions, clinicians or therapists provide briefings about disorders and illnesses in everyday language to make this knowledge easily accessible to patients and their relatives. Insights about medical and biochemical, but also psychosocial factors of illnesses can help patients and their families to develop a fundamental understanding of disease and putative therapies (Anderson, Hogarty, & Reiss 1980). It has been described that patients and their social environment feel more relaxed, less ashamed and guilty, and much more in control of the condition by having a greater level of understanding concerning the multifactorial causes and trajectory of their illness (Lucksted et al. 2012; Schiffman et al. 2015). For this approach to be helpful in coping with illness or to improve mental hygiene in patients, it is particularly useful to disentangle the intricate complexities of a disease as outlined in the biopsychosocial model of health. A strong focus should be given to the effects of stress, how the human body copes with stressful life events, and which factors (biomedical or others) make us particularly vulnerable to disease (Sapolsky 2017). It is thus of paramount importance that psychoeducation is informed by the biopsychosocial model of health rather than the traditional medical model of disease since the former strongly empowers patients and facilitates self-efficacy. In effect, psychoeducation paves the way for partially handing over the responsibility to the patient, thereby fostering a commitment to the therapy, lowering relapse rates, and ensuring long-term involvement.

The case of K.

K., female, 17 years old

K. is the middle child of three siblings. Her parents are divorced. At the time of the first session, she lived alone in a small apartment. Although her grades were satisfactory, K. left school. After leaving school, she started working with a manufacturer of mechatronics systems as a trainee. She has little contact with parents, siblings, or friends. Her main social interactions are with A., a social worker who checks on her a few times per week. K. comes across as intelligent, empathic, curious, and friendly. She has been regularly visiting a psychiatrist who prescribed her antidepressants (selective serotonin reuptake inhibitors (SSRI)). She has also been seeing a psychotherapist for several months. Recently, she has missed a few psychotherapy sessions.

During the first session after the hiatus, she mentions that she has not been to work for three weeks and feels guilty about not going to work. She then admits that she also has not left the house in weeks, with the only exception of leaving the apartment for this therapy session. In the course of the conversation, she notes in passing that she has stopped taking the medication prescribed by her psychiatrist. She does not see the need to take antidepressants because she does not feel ill, but only feels she lacks a purpose in life. In her opinion, she is going through a phase in which she does not want to meet anybody due to negative experiences related to her family, especially her mother and older sister – a family, she feels, that she does not belong to. She describes herself as inadequate to her mother's expectations and overwhelmed by life in general. Going to work, doing household chores, and even following daily routines, like taking a shower, preparing food, and getting out of bed, feel like unmanageable tasks at times.

She admits that she has decided to come to the therapy session because she was curious about the psychoeducation part. She was promised by the psychotherapist that a neuroscientist/philosopher of neuroscience (I.S-J) will be present to give her some general insights into the psychopathology of depression and how mental illnesses can be entangled with environmental and social factors. Based on the questions K. asked and answers she gave during this session as well as in previous therapy sessions, it could be assumed that she is indeed interested in the issue and cognitively receptive to benefiting from a psychoeducational intervention.

At the beginning of the session, the psychotherapist encourages K. to describe her day and the things she likes to do. K. reports as asked, but soon shifts the focus of her story and addresses the feelings she experiences when she faces the challenges of daily routines. Although she emphasizes the advantages of not leaving the apartment or not communicating with anybody, she quickly turns to the negative aspects that her behavior entails, including feelings of sadness, purposelessness, and an increase in anxiety the longer the situation persists. The psychotherapist includes me in the conversation by asking whether the symptoms K. is experiencing may have something to do with her discontinuing the medication prescribed by her psychiatrist. I start describing the effects that SSRIs usually have and why patients with depression can benefit from taking them. I emphasize that the information I am giving is general and that I cannot comment on any specifics of K.'s case. I also point out unwanted effects which patients may experience due to the medication. The psychotherapist asks K. which things she did when taking the antidepressants – activities that may or may not have had anything to do with the fact that she was on medication – leaving it open to K. to interpret what is said.

K. recounts activities from several weeks ago, mentioning that she actually liked going to work and attending group sessions that A., the social worker, organized. She recalls that she met her father and she enjoyed drawing pictures of fantasy characters. She also reports that she managed to keep on top of household chores as they were actually not that burdensome given the small size of the apartment. K. wants to know whether she can manage to do these things without taking antidepressants. I tell her about serotonin activity (that is targeted by the SSRI), which becomes dysregulated during episodes of stress due to feedback mechanisms of stress hormones acting on certain brain areas. I mention other therapeutic measures or non-therapeutic activities that have been shown to correlate with increased serotonin activity, albeit in a modest manner.

K. and the psychotherapist discuss triggers that K. perceives as stressful and which methods she can apply when experiencing stressful situations. The therapist invites me to reflect on the biopsychosocial model, especially on how socially stressful situations can modulate physiological responses. I resort to examples from animal studies emphasizing that these are by no means indicative of what happens in humans in particular situations. By way of illustration, I point out that humans have the advantage of being able to exploit excellent sociocultural means that allow stress regulation and that can contribute to feeling more relaxed and regaining energy. The therapist leads K. back to discussing things that she enjoys and asks which of these things would be worth pursuing if she could change her current habits. K. mentions two things she misses because of not leaving her apartment. The first one is going for an ice cream and the second one is spending more time with A., her social worker (who is now only coming briefly to K.'s apartment to check if everything is OK).

At the end of the session, the therapist and K. arrange another appointment, while K. asks whether she can have further information about the biopsychosocial aspects of depression by means of psychoeducation. I agree to join the next session.

The next session takes place two weeks later. K. is accompanied by A., the social worker. K. says she feels better, can do household chores again, has been looking after herself, and has left the flat a few times to go shopping and to meet her father. She still has not been back to work. The therapist does not address her comment about work, but instead compliments her about the achievements. K. mentions that after visiting her psychiatrist, recently, she has started to take antidepressants again. She says that she is determined to take them until she feels well enough to master her everyday life without drugs.

In this session, psychoeducation plays a much smaller role. K.'s main intention is to ask me to repeat things I mentioned to her in the previous session, so that A. is on the same page as K. with respect to the biopsychosocial factors influencing depression. The psychotherapist and K. discuss (personal and interpersonal) resources that K. can rely on to tackle everyday challenges. K. emphasizes A.'s role in the healing process. The language of the whole session is very collaborative. K. reflects on triggers that cause strongly adverse feelings, including some topics that came up during a telephone conversation with her mother. She interprets these triggers in light of family tensions between her parents and between her and her siblings. She mentions the liberation as well as the fear of making her own decisions since not living at home anymore. K. ponders how she wants to react to such stressful events in the future and what she may do to avoid unnecessary stressors. She outlines a story that illustrates how she wants to be perceived by her mother and other family members. The narrative focuses on her current life choices with short explanatory remarks concerning her personal history. She also very briefly ventures into possibilities that may open up in the future, especially in relation to a potential profession, but then she quickly terminates these ideas.

In the debriefing after the session, the therapist points out that K. has reflected on her condition as being an understandable physiological reaction rather than a unique malady that could only befall her because of a certain deficit. This alternative take on the situation has led to K. feeling significantly less guilty and ashamed about her condition than before.

Limitations of the narrative approaches

Despite the clear advantages of narrative approaches, this method has significant limitations when dealing with severe cases of psychopathologies, such as autism, extreme aggression, severe bipolar disorder, overwhelming anxiety, severe obsessive symptoms, and psychosis (White 1995). In many such cases, narrative approaches can only be employed once psychiatric medication becomes effective. Albeit not always, but in many cases (as in the case described above) medication can open up opportunities to begin a patient's story development. The therapist can then work to enrich the story within this paradigm. It is, however, important to continue exploring narrative approaches despite medication (i.e, medication should not be the therapeutic end point following a diagnosis).

Another limitation that psychiatrists and therapists might face when using a narrative approach occurs in cases in which patients have an intellectual disability or significant cognitive impairments. These patients will probably not benefit as much from narrative therapies as patients with greater cognitive function. In addition, some patients with "normal" cognitive function may simply prefer a diagnosis that is largely described in biomedical terms and will consciously or subconsciously resist narrative approaches as too "vague" and "unscientific." In these cases, it won't be helpful to persuade patients since the narrative approach requires patients to collaborate and make a significant effort in the co-authoring process.

Conclusion

Since healthcare, especially mental healthcare has experienced a re-thinking of its underlying conceptual framework, slowly shifting from the medical model of disease to a biopsychosocial model of health, narrative approaches have become increasingly important in clinical practice. By aiming at integrating sociocultural and psychosocial aspects in how disorders are conceptualized, traditional medical-oriented processes – interviewing, diagnosing, and treatment – have become enriched with narrative tools. Narrative approaches have been shown to be powerful tools to foster patients' ability to cope with physical or mental illnesses. Yet, the success seems to be tightly linked to the physician's or therapist's ability to influence and modify the patient's narrative into a story that resonates with the patient's own life stories. A fundamental requirement for this co-construction and co-creation process is the patient's understanding of the disorder itself and the multiple factors that can influence it. To make a diagnosis accessible and subject to transformation, clinicians or psychotherapists need to translate scientific and medical knowledge about the disease into everyday language using psychoeducational methods. Critical to this endeavor is, however, the healthcare professional's own concept of disease. To co-construct empowering stories about healing, narrative clinicians and therapists must strongly commit to a biopsychosocial model and acknowledge the shortcomings of reductionist interpretations of disorders – a conceptual shift that has yet to be mirrored in medical curricula.

References

Anderson C M, Hogarty G E, Reiss D J (1980) Family Treatment of Adult Schizophrenic Patients: A Psycho-educational Approach. *Schizophrenia Bulletin* 6.3: 490–505.

Benedetti F (2013) Placebo and the New Physiology of the Doctor-Patient Relationship. *Physiological Reviews* 93.3: 1207–1246.

von Bertalanffy L (1972) Zu einer allgemeinen Systemlehre. In *Organisation als System*. *Organization und Führung*, reprinted by Knut Bleicher, 29–46. Wiesbaden: Gabler Verlag.

Boorse C (1975) On the Distinction between Disease and Illness. *Philosophy and Public Affairs* 5.1: 49–68.

Boorse C (1977) Health as a Theoretical Concept. *Philosophy of Science.* 44.4: 542–573.

Carrey N (2007) Practicing Psychiatry Through a Narrative Lens. In *Narrative Therapy: Making Meaning, Making Lives*, ed. Catrina Brown, Tod Augusta-Scott, 77–101. London: SAGE Publications.

Chipidza F E, Wallwork R S, Stern T A (2015) Impact of the Doctor-Patient Relationship. *The Primary Care Companion for CNS Disorders* 17.5: 10.4088/PCC.15f01840.

Deacon B J (2013) he Biomedical Model of Mental Disorder: A Critical Analysis of Its Validity, Utility, and Effects on Psychotherapy Research. *Clinical Psychological Review* 33.7: 846–861.

Dennett D C (1992) The Self as a Center of Narrative Gravity. In *Self and Consciousness: Multiple Perspectives*, ed. Frank S. Kessel, Pamela M. Cole, Dale L. Johnson, and Milton D. Hakel, 103–115. Hillsdale: Lawrence Erlbaum.

Engel G L (1977) The Need for a New Medical Model: A Challenge for Biomedicine. *Science* 196.4286: 129–136.

Fabrega H Jr. (1975) The Need for an Ethnomedical Science. *Science* 189.4207: 969–975.

Gazzaniga M S (1992) The Self as a Center of Narrative Gravity. In *Self and Consciousness: Multiple Perspectives*, ed. Frank S. Kessel, Pamela M. Cole, Dale L. Johnson, and Milton D. Hakel, 6–88. Hillsdale: Lawrence Erlbaum.

Gomez E A (1996) Radical Constructivism and Narrative. *Journal of American Academy of Psychoanalysis* 24.1: 1–14.

Gurova L (2013) 'Understanding It Makes It Normal': Is It a Reasoning Fallacy or not? *Journal of Evaluation in Clinical Practice* 19.3: 524–527.

Hacking I (1995) The Looping Effects of Human Kinds. In *Causal Cognition*, ed. Sperber, Dan, Premack, David, and Ann J. Premack. Oxford: Oxford University Press.

Kirmayer L J, Crafa D (2014) What Kind of Science for Psychiatry? *Frontiers in Human Neuroscience* 8.e27407: 1–12.

Laing R D (1971) *The Politics of the Family and Other Essays*. London: Routledge.

Launer J (1999) A Narrative Approach to Mental Health in General Practice. *British Medical Journal*, 318.7176: 117–119.

Lucksted A, McFarlane, W, Downing D, Dixon L (2012) Recent Developments in Family Psychoeducation as an Evidence-Based Practice. *Journal of Marital & Family Therapy* 38.1: 101–121.

Ludwig A M (1975) The Psychiatrist as Physician. *Journal of the American Medical Association* 234.6: 603–604.

Moncrieff J (2008) *The Myth of the Chemical Cure: A Critique of Psychiatric Drug Treatment*. Basingstoke: Palgrave Macmillan.

Moncrieff J, Cohen D (2005) Rethinking models of psychotropic drug action. *Psychotherapy and Psychosomatics* 74.3: 145–153.

Nylund D, Tilsen J (2006) Pedagogy and Praxis: Postmodern Spirit in the Classroom. *Journal of Systemic Therapies* 25.4: 21–31.

Parry A, Doan R E (1994) *Story revisions: Narrative Therapy in the Postmodern World.* New York: Guilford Press.

Sapolsky R M (2017) *Behave: The Biology of Humans at Our Best and Worst*. New York: Penguin Press.

Sarto-Jackson I (2018) Time for a Change: Topical Amendments to the Medical Model of Disease. *Biological Theory* 13.1: 29–39.

Schiffman J, Reeves G M, Kline E, Medoff D R, Lucksted A, Hoagwood K, Fang L J, Dixon L B (2015) Outcomes of a Family Peer Education Program for Families of Youth and Adults with Mental Illness. *International Journal of Mental Health* 44.4: 303–315.

Simblett G (1997) Leila and the Tiger: Narrative Approaches to Psychiatry. In *Narrative Therapy in Practice: The Archeology of Hope*, ed. Monk, Gerald, Winslade, John, Crocket, Kathie, and David Epston, 132–145. San Francisco: Jossey-Bass.

White M (1995) Psychotic experiences and Discourse. In *Re-authoring Lives: Interviews and Essays*, 112–153. Adelaide: Dulwich Centre Publications.

White M, Epston D (1990) *Narrative Means to Therapeutic Ends*. New York: Norton.

Winslade J, Monk G (2000) *Narrative Mediation: A New Approach to Conflict Resolution*. San Francisco: Jossey-Bass.

THREE ROLES OF NARRATIVES IN THE TREATMENT OF CHRONIC PAIN*

Nina A. Atanasova

Department of Philosophy and Religious Studies, The University of Toledo, Toledo, Spain
email: nina.atanasova@utoledo.edu

Abstract:
In this paper, I discuss the roles narratives play in the diagnostics, treatment, and recovery of chronic pain patients. I show that the successes of this narrative approach to the treatment of chronic pain support the biopsychosocial model of disease. The central example of narrative interventions discussed in the paper is pain neuroscience education. This is an intervention which aims at helping chronic pain patients reconceptualize their pain experiences so as to align them with neuroscientific knowledge of pain. Multiple clinical trials have established the success of these interventions in pain reduction. This shows that neuroscience pain education is in fact an evidence-based approach. I conclude that narrative and evidence-based medicine are compatible.

Key words: Neuroscience Pain Education, Narrative Medicine, Biopsychosocial Model of Disease, Chronic Pain.

Introduction

Chronic pain has traditionally been conceptualized through the medical model of disease and has been treated correspondingly, through surgery, medication, and physiotherapy. However, these interventions alone are too often unsuccessful in the treatment of intractable pain (Goesling et al. 2018). Complementary biopsychosocial approaches, including narrative interventions, have proven successful in the reduction of chronic pain patients' suffering (Louw et al. 2016; Watson et al. 2019).

The biopsychosocial model of disease adopted in narrative medicine is often opposed to the medical model characteristic of evidence-based medicine (Ghaemi 2009; Marini 2016). I will argue that this opposition is specious. Contemporary narrative-based approaches to the treatment of chronic pain have already generated significant evidence that the integration of psychosocial and biomedical approaches can be successful. I take that to show that narrative-based medicine and evidence-based medicine are compatible and that the biopsychosocial model is an extension of the biomedical model rather than its competitor.

In what follows, I will introduce the mainstream conceptualization of pain in contemporary neuroscience and distinguish between adaptive and maladaptive pain. I will then present examples of three roles narratives play in the treatment of pathological chronic pain. To conclude, I will discuss the implications of the examples of narrative interventions discussed for the compatibility between narrative and evidence-based medicine.

1. What is pain?

The International Association for the Study of Pain defines pain as: "An unpleasant sensory and emotional experience associated with, or resembling that associated with, actual or potential tissue dam-

* This paper was inspired by the discussion of the role of narratives in psychiatry which took place in 2019 at the biennial meeting of the ISHPSSB in Oslo. I am indebted to my fellow symposium presenters, Isabella Sarto-Jackson, Lilia Gurova, and Paola Hernandez-Chavez as well as to our engaged audience.

age."[1] Correspondingly, contemporary neurobiological models of pain recognize three dimensions of the experience: (1) nociceptive (sensory experience); (2) affective (unpleasant experience); and (3) cognitive (reflection on the experience). The different modalities have been shown capable of dissociation from one another through various interventions such as analgesics, narcotics, and hypnosis. For example, Rainville et al. (1999) used hypnosis to dissociate the sensory and affective dimensions of pain. They showed that hypnotic suggestion can decrease unpleasantness with no change to the perceived intensity of the experimentally induced pain and, conversely, that under hypnotic suggestion the modulation of the sensory dimension may leave the affective evaluation of pain unpleasantness unaltered. Narcotics such as morphine have been shown to eliminate the unpleasantness of pain without affecting the awareness of it (Barber 1959). That is, the affective dimension may be altered without change in sensory or cognitive processing of the painful experience. Finally, affective and cognitive dimensions of pain may be experienced in the absence of nociceptive stimulation. This is quite obvious in phantom limb patients who report pain localized in amputated limbs.

Contemporary biomedical conceptualizations of sensory pain are grounded in the so-called *gate control theory* introduced by Melzack and Wall (1965). They proposed a model in which stimulation of the nociceptors (receptors of noxious stimuli located on the skin, muscles, joints, and other parts of the body where tissue damage can occur) produces neuronal signals which travel through two types of neuronal fibers (A-delta and C) to the spinal cord. From there, the signals may be further projected to the brain where they result in the experience of pain. However, the site at which the nerve fibers enter the spinal cord contains a 'gate system' which is controlled by a complex mechanism. When tissue damage is inflicted, the A-delta fibers send strong impulses which open the gate, and the information about the damage is quickly sent to the brain. This produces the experience of sharp pain, which motivates the retraction from the source of injury. If the larger non-nociceptive A-beta fibers are stimulated at the same time (for example by rubbing the sight of injury) they send impulses to the gate which close it. This is how pain experience can be modulated. However, with prolonged activity of the smaller A-delta and C fibers, the larger fibers may adapt, and the gate will open again. The smallest and slowest C fibers are typically responsible for prolonged signals. This results in the experience of throbbing pain which motivates nursing of wounds. The opening and closing of the gate may further be modulated by central activities such as fear and anxiety. Based on evidence for such modulation, Melzack and Wall further posit the existence of a mechanism, which they term 'central control trigger,' that is responsible for selective brain processes that modulate sensory input (p. 976).

Melzack and Casey (1968) further elaborate on this model by stressing that pain cannot be reduced to a purely sensory experience. It is a complex phenomenon which, in addition to the sensory dimension, also has affective-motivational and cognitive dimensions. Melzack and Casey maintain that all three dimensions need to be accounted for in the conceptualization of the behaviors which characterize pain (p. 434). An important implication of this view is that pain can be modulated through interventions on any of the three constitutive dimensions. They conclude that "Pain can be treated not only by trying to cut down the sensory input by anesthetic, surgical intervention and the like, but also by influencing the motivational-affective and cognitive factors as well" (p. 435). The multidimensionality of pain has since been corroborated in a number of experimental studies, particularly studies of dissociation (Moayedi & Davis 2013).

2. Chronic Pain as a Disease

Pain is a natural part of human existence. It has survival value. It helps us avoid threats to our lives and wellbeing. It is an evolutionary adaptation which was selected for through natural selection. Its adaptational advantage is especially clear when put in perspective against the misfortune of sufferers from congenital insensitivity to pain (CIP). People with this condition lack pain sensations regardless of their otherwise intact sensory system. They tend to live very short lives due to multiple injuries they sustain as children. Common injuries include, among others, severe arthritis and self-mutilated fingers resulting from chewing on them (e.g., Nabiyev et al. 2016; Nagasako et al. 2003). CIP sufferers literally wear out their bodies because they do not notice the severe injuries they inflict on themselves. This condition proves that acute pain is actually good (Emad et. Al 2007).

[1] Special thanks go to my student, Alex Hollatz-Guastella, for pointing out the recent update of the IASP definition of pain, as of July 2020.

However, not all pain is good. Phantom limb pain, low back pain, fibromyalgia, and other pain disorders have no survival value. They are maladaptive. Pain experiences that persist for more than three months in the absence of tissue damage, even if they are residual from a past injury, are classified as chronic pain and are considered diseases. Throughout the twentieth century, pain was conceptualized through the medical model of disease and was treated accordingly. Within this model, medical doctors, nurses, and physical therapists treat organic disorders through physical interventions such as surgery, medication, physiotherapy, etc. The field of pain studies only started to take shape in the 1960s (Collier 2017). The comorbidity of chronic pain and mental illness, such as anxiety and depression, was not fully appreciated until very recently (Goesling et al. 2018).

> Historically, the bio-medical treatment model of chronic pain focused primarily on relieving the physical symptoms. For example, a person with low back pain would be treated with surgical interventions, injections, and/or medication. Yet, less than half of patients improve following most surgical interventions and medications, especially opioids (Goesling et al. 2018: 12).

Under the diagnostic assumptions of the medical model, chronic pain is often intractable and patients have historically been accused of malingering (Collier 2017). However, recent developments in neuroimaging technology have proven utility in discerning whether a patient does indeed have a painful experience in the lack of tissue damage. This is the case with conditions such as hyperalgesia, allodynia, and fibromyalgia, for example (Hardcastle & Stewart 2009). Additionally, the biopsychosocial model of disease has been very successful for the study and treatment of chronic pain. Multidisciplinary approaches based on neurobiology as well as psychology and the social sciences have been utilized in the treatment of chronic pain patients. In this form of comprehensive medical care, conversation and storytelling play key roles. I will explore them in the following section.

3. Three Roles of Narratives in the Treatment of Chronic Pain

Within the biopsychosocial model of chronic pain, it has been recognized that pain experiences require interpretation. One needs to be able to identify what the painful experience means. Is it an indication of tissue damage as in acute pain or is it a pathology which itself needs to be treated? The answer to this question is often found in the story about the experience of the person in pain. When the pain turns out to be a pathology, the patient will need to tell this story to the physician who, in turn, will identify the likely diagnosis and tell a story that explains what happens to the patient physiologically while they are having this experience. Finally, a crucial component of the patient's recovery will involve the reconceptualization of the painful experience. The patient will tell another story, a story of healing. At the different stages of this process, narratives play different roles. I will discuss three of them: diagnostics, treatment, and recovery.

3.1. Diagnostics: Narrative medicine

Chronic pain is often without easily identifiable organic causes. It may result from an injury but then persist long after the damaged tissue has healed. When all available physical diagnostic techniques have failed, patients may face distrust from others with whom they are in personal or professional relationships – sometimes even healthcare providers (Collier 2017). This is why a narrative approach to diagnosing chronic pain is crucial (Charon 2005; Loeser 2005). Patients need to be able to tell their story and to be heard by their physicians. This enables comprehensive diagnostics. When patients tell their stories rather than merely reporting their symptoms, they mention various experiences which provide psychological and social context to the pain symptomatology. Charon (2005) discusses the case of posttraumatic stress disorder as a common cause of chronic pain. Allowing patients to express an uninterrupted narrative provides the opportunity to trace the onset of their experiences to that past trauma.

Watson et al. (2019) provide empirical evidence suggesting that patients' narrations during the diagnostic stage of chronic pain improve their treatment outcomes. Letting patients tell their own stories helps them feel heard and fosters a personalized approach to their treatment (p. 1140.e17). These findings are consistent with the multidimensional conceptualization of pain. Dysfunctional emotional and cognitive processes may be responsible for the experiences of pain in the absence of tissue damage or other organic cause. Such processes may be identified through the patient's narrative and could then be treated with interventions which target and alter the dysfunctional emotional and cognitive processes.

3.2. Treatment: Pain Neuroscience Education

Pain neuroscience education (PNE) is an intervention targeting cognitive reappraisal of chronic pain. It was inspired by Gifford and Butler's (1997) call for integration of the science of pain within clinical practice. They advocated that clinicians and physiotherapists should align their approaches to treatments of pain with gate control theory, which was already well-established in neuroscience, along with the multidimensionality of pain experience. Gifford (1998) further developed this proposal by providing an argument for the necessity of educating patients about the mechanisms of pain within the interventions employed for their treatment.

His view is grounded in evidence that the experience of pain has affective and cognitive as well as sensory dimensions. Patients' prior experiences and their cultural contexts inform their beliefs about, and affect their emotional attitudes toward, pain. These beliefs and emotional attitudes are instantiated in the central nervous system and are part of the mechanisms of pain processing. Emotions such as fear and anxiety are adaptive when they motivate a healthy individual to avoid injury by avoiding dangerous activities and environments. However, in chronic pain patients, the functioning of the pain processing neural circuitry shifts from processing of nociceptive stimuli to more widespread neuronal activity. This leads to experiences of pain which are underdetermined by sensory inputs and are dominated by maladaptive cognitive and affective processing. Individuals who hold such maladaptive beliefs, Gifford points out, may be overfocusing on their thoughts and thus keeping the control gate in their pain processing system open. This, in turn, produces the maladaptive experience characteristic of chronic pain. Gifford's central proposal is that altering the maladaptive beliefs of patients about the physiology of their pain will facilitate their recovery (p. 33).

Gifford's work has since inspired an approach to the clinical treatment of pain which is based on patient education about the physiology and neuroscience of pain. The approach is based on narratives, using metaphors and storytelling to educate patients about their pain with the goal of aligning their beliefs with scientific knowledge about pain and ultimately contributing to their recovery. This approach is different from other psychosocial interventions that aim at changing the beliefs and behaviors of patients, such as cognitive behavioral theory (CPT) (Moseley & Butler 2015). While CBT is based on advising patients on how to cope with pain, PNE aims at changing fundamental beliefs, thereby intervening in the very mechanism which produces the experience of pain. Rather than managing pain, it aims at reducing it. The method is mediated by narratives, and a number of studies have shown its effectiveness in reducing the cognitive and affective experience of pain (Louw et al. 2016; Watson et al. 2019).

3.3. Recovery: Pain Reconceptualization

The multidimensional model of pain recognizes three stages of pain processing (Wade et al. 2011). The first stage, according to Wade et al., is characterized by nociception, which is present in acute pain but may be absent in chronic pain. The second stage involves minimal cognitive processing and concerns the individual's immediate responses to the painful sensation. The third stage is characterized by self-reflection and complex cognitive processing of the long-term implications that pain has for one's life. Complex emotions such as feelings of depression, anxiety, frustration, anger, and fear may dominate the patient's experience at this stage. The chronic experience of these pain-induced emotions depends on the degree of catastrophic thinking. Catastrophizing may include excessive focus on pain sensations (rumination), exaggeration of the threat (magnification), and feelings of helplessness (p. 314). Wade et al. further cite neuroimaging studies which indicate that chronic pain patients show greater activity in all pain processing brain areas, including the sensory cortex, compared to controls, in response to identical pain inducing interventions (p. 318). Such experimental results suggest that psychopharmacological and psychosocial interventions may be successful at reducing pain by targeting the neural mechanisms responsible for the third stage of pain processing.

Indeed, non-opioid pharmacological treatments such as antidepressants have been shown to produce moderate analgesic effects (Goesling et al. 2018, p. 4). A psychosocial intervention that has shown moderate positive results for managing chronic pain is CBT. Finally, according to PNE proponents, it is even more successful than CBT as it intervenes to reduce pain rather than simply manage it (Mosley & Butler 2015). The main difference between CBT and PNE is that the former offers advice on behavior whereas the latter uses storytelling in order to reconceptualize the experience of pain. PNE aims at changing the narratives that chronic pain patients use in order to make sense of their experiences.

The role of narrative in recovery has also been promoted in alternative medicine. Narrative is central to complementary interventions which similarly rely on reconceptualization of pain and self-construction. Some of them rely on sharing one's pain narrative in support groups and creative writing (Hovey et al. 2018), others on constructing a narrative of spiritual growth and self-discovery (Belton 2019). The success of such interventions is anecdotal. However, one can speculate that the relief chronic pain patients achieve through creativity and spirituality is mediated by changes in catastrophizing. In this way, narratives are therapeutic.

4. The Biopsychosocial Model of Pain: An Evidence-Based Approach

The three roles of narratives in the care for chronic pain patients and their therapeutic success support the biopsychosocial model of pain. Having said that, I need to quickly qualify this statement. The biopsychosocial model was conceived as a 'holistic', nonreductive model of disease (Engel 1977; Bolton & Gillett 2019). The model postulates the reality of psychological and social causes of disease, which cannot be reduced to molecular biology. Engel introduced the model as a critique of the mainstream medical model of disease which reduced it to molecular biology. Simultaneously, Engel argued, medicine assumes the mind-body dualism in which the psychosocial is separate from the biophysical and where only the biophysical can have a place in medicine. He makes the case that psychosocial factors contribute to organic disease and biological factors contribute to psychiatric diseases. This, according to Engel, is why the methods of biomedicine and psychiatry must be adopted in the treatment of all diseases.

While I agree with this conclusion, I am not convinced that it is incompatible with physicalist reductionism. On the contrary, I see the model's value in providing a reductive account of how psychosocial interventions are ontologically grounded in biology. PNE integrates, rather than uses in parallel, knowledge from neuroscience, psychology, and educational science. This is what makes PNE so successful. Its psychosocial interventions are simultaneously neurological interventions. The success of these psychosocial interventions is explained by their dependence on the neural basis that instantiates them. This makes the biopsychosocial model of chronic pain reductive – there is no room in it for mysterious nonphysical causation, and interventions are ultimately physical interventions. The biopsychosocial model of chronic pain is in this sense an extension of the medical model as opposed to its rival.

The successful integration of psychology, neuroscience, and educational science in the case of chronic pain may be indicative of the possibility of integrating seemingly exclusive approaches for other medical conditions as well. It may be that the medical model needs to be refined in order to provide integrated comprehensive healthcare rather than be replaced. Charon (2001), and many others since, advocate narrative medicine as providing a more successful approach to healthcare than symptom-centered approaches. She identifies multiple roles that narratives play in the context of patient care, caregiver self-care, and relationships between healthcare providers as well as their interactions within society in general. All of these domains need to be fostered for the success of medicine, Charon recommends. Narrative medicine is the fulfilment of the biopsychosocial approach to disease and health. The narrative-based approach to the treatment of chronic pain has already proven its success (Louw et al. 2016; Watson et al. 2019). It is now also an evidence-based approach. Perhaps narrative-based approaches in other areas of medicine will eventually mature to the extent that they become evidence-based as well. In fact, this trend has already been discussed in Marini (2016).

Conclusion

The examples discussed above show the importance of narratives in the conceptualization of conditions which have traditionally been treated with medical interventions alone. The effectiveness of narrative interventions in chronic pain treatment proves that the biopsychosocial model can be integrated with the medical model. The systematic successes of narrative-based therapeutic interventions discussed above also show that narrative-based medicine is compatible with evidence-based medicine.

References

Barber T X (1959) Toward a Theory of Pain: Relief of Chronic Pain by Prefrontal Leucotomy, Opiates, Placebos, and Hypnosis. *Psychological Bulletin* 56(6): 430–460.

Belton J (2019) Exploring the Meanings of Pain: My Pain Story. In *Meanings of Pain, Volume 2: Common Types of Pain and Language,* ed. van Rysewyk, 1–15. Springer Nature Switzerland.

Bolton D, Gillett G (2019) *The Biopsychosocial Model of Health and Disease: New Philosophical and Scientific Developments.* Palgrave Macmillan.

Charon R (2001) Narrative Medicine: A Model for Empathy, Profession, and Trust. *Journal of American Medical Association* 15: 1897–1902.

Charon R (2005) A Narrative Medicine for Pain. In *Narrative, Pain, and Suffering. Progress in Pain Research and Management,* ed. D. B. Carr, J. D. Loeser, D. B. Morris, 34: 29-44. Seattle: IASP Press.

Collier R (2017) A short history of pain management. *CMAJ* 8(190): E26–27.
doi: 10.1503/cmaj.109-5523

Emad Y, El Yasaki A, Ragab Y, Khalifa M, Moawayh O, Salama M (2007) Arthritis in a child secondary to congenital insensitivity to pain and self-aggression. Why and when pain is good? *Clinical Rheumatology,* 26(7), 1164–1166.

Engel G L (1977) A Need for a New Medical Model: A Challenge for Biomedicine. *Science* 196(4286): 129–136.

Ghaemi S N (2009) The rise and fall of the biopsychosocial model. *The British Journal of Psychiatry* 195:3–4.

Goesling J, Lin L A, Clauw D J (2018) Psychiatry and Pain Management: at the Intersection between Chronic Pain and Mental Illness. *Current Psychiatry Reports* 20:12.

Gifford, L. (1998). "Pain, the Tissues and the Nervous System: A conceptual model." *Physiotherapy 84*(1): 27-36.

Gifford L S, Butler D A (1997) The Integration of Pain Science into Clinical Practice. *Journal of Hand Therapy* 10(2): 86–95.

Hardcastle V G, Stewart C M (2009) fMRI: A Modern Cerebrascope? The Case of Pain. In *The Oxford Handbook of Philosophy and Neuroscience* ed. John Bickle, 179–199. Oxford: Oxford University Press.

Hovey R B , Khayat V C, Feig E (2018) Cathartic Poetry: Healing through Narrative. *The Permanente Journal* 22: 17–196.

Loeser J D (2005) Pain, Suffering, and the Brain: a Narrative of Meanings. In *Narrative, Pain, and Suffering. Progress in Pain Research and Management,* ed. D. B. Carr, J. D. Loeser, D. B. Morris, 34: 17–27, Seattle: IASP Press.

Louw A, Zimney K, Puentedura E J, Diener I (2016) The efficacy of pain neuroscience education on musculoskeletal pain: A systematic review of the literature. *Physiotherapy Theory and Practice* 32(5): 332–355.

Marini M G (2016) *Narrative Medicine.* Springer Cham.

Melzack R, Casey K L (1968). Sensory, Motivational, and Central Control Determinants of Pain. In *The Skin Senses. Proceedings of the First International Symposium on the Skin Senses, ed.* D. R. Kenshalo, : 423-439. Springfield, Illinois: Charles C. Thomas Publisher.

Melzack R, Wall P D (1965) *Science* 150(3695): 971–979.

Moayedi M, Davis K D (2013) Theories of pain: from specificity to gate control. *Journal of Neurophysiology* 109: 5–12.

Moseley G L, Butler D S (2015) Fifteen Years of Explaining Pain: The Past, Present, and Future. *The Journal of Pain* 16(9): 807–813.

Nabiyev V, Kara A, Aksoy M C (2016) Multidisciplinary assessment of congenital insensitivity to pain syndrome. *Child's Nervous System* 32(9), 1741–1744.

Nagasako E M, Oaklander A L, Dworkin R H (2003) Congenital insensitivity to pain: an update. *Pain* 101(3): 213–219.

Rainville P, Carrier B, Hofbauer R K, Bushnell M C, Duncan G H (1999) Dissociation of sensory and affective dimensions of pain using hypnotic modulation. *Pain* 82(2): 159–171.

Wade J B, Riddle D L, Price D D, Dumenci L (2011) Role of pain catastrophizing during pain processing in a cohort of patients with chronic and severe arthritic pain. *Pain* 152: 314–319.

Watson J A, Ryan C G, Cooper L, Ellington D, Whittle R, Lavender M, Dixon J, Atkinson G, Cooper K, Martin D J (2019) Pain Neuroscience Education for Adults With Chronic Musculoskeletal Pain: A Mixed-Methods Systematic Review and Meta-Analysis. *The Journal of Pain* 20(10): 1140.e1–1140.e22.

FEIGNED NARRATIVES DO NOT ALWAYS SATISFY NEEDS: THE CASE OF FACTITIOUS DISORDERS*

Lilia Gurova

Department of Cognitive Science and Psychology,
New Bulgarian University, Sofia, Bulgaria
email: lgurova@nbu.bg

Abstract:

When Bradley Lewis announced in 2014 that psychiatry needed to make a "narrative turn", he backed up his appeal as follows: (1) the different explanatory models of mental disorders that are currently competing in psychiatry tell us different stories about mental health; (2) none of these stories has the privilege of being the only true one, and its alternatives the wrong ones; (3) the choice of a model in each case should be made in dialogue with the patient in order to ensure that the model will be chosen that best meets the patient's goals and desires and, accordingly, would best support the process of recovery. The latter suggestion, however, is not easy to follow when the patients' subjective goals and desires diverge from the clinical goal of returning the patients to a normal way of life, as is the case with the so-called factitious disorders. The problem is worsened by the theory-ladenness of the interpretations of patients' first-person narratives. This paper argues against a common assumption that biases our understanding of abnormal behavior, in particular the behavior of those who feign illness. The assumption in question is the following: that such behavior satisfies certain – possibly unknown – psychological needs.

Key words: understanding mental disorders, narrative turn in psychiatry, theory-ladenness of understanding, first-person narratives, factitious disorders, self-engaging behavior.

> "We also tell stories just because we cannot stop, because they fascinate and engage us even if we know they are untrue."
>
> (Boyd 2009, p. 1)

1. Introduction: the call for a narrative turn in psychiatry and the focus on patients' first-person narratives

On April 29, 2013, just weeks before the official release of the fifth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5), Thomas Insel, who was then director of the National Institute of Mental Health (NIMH), expressed in a blog post the intent of NIMH to "re-orient its research away from DSM categories" (Insel 2013). Admitting the usefulness of DSM (of all of its editions) as "dictionary" ensuring that the same terms are used in the same way by all professionals, Insel at the same time stressed the most

* An early version of a part of this text was presented at the ISHPSSB symposium "Reassessing the role of narratives in psychiatry" (Oslo, 7–12 July 2019). I am indebted to Nina Atanasova for the organization of this symposium and for the invitation to take part in it. I also thank Paola Hernández-Chávez and Isabella Sarto-Jackson for the memorable discussions on the main shortcomings of the medical model of mental disorders. These discussions strengthened my conviction that, although the alternatives to the medical model are still in infancy or underdeveloped, it makes sense to put effort into them. The present paper, I hope, is a step in this direction.

serious defect of the DSM categories – their lack of validity due to the fact that these categories refer to "clusters of clinical symptoms, not any objective laboratory measure.""Patients with mental disorders deserve better," said Insel. That is why NIMH had started working on a project which aims to "transform diagnosis by incorporating genetics, imaging, cognitive science, and other levels of information to lay the foundation for a new classification system" (Insel 2013). This project, which is known by the name "Research Domain Criteria" (RDoC), is based on several assumptions but the most important among them is that "mental disorders are biological disorders involving brain circuits that implicate specific domains of cognition, emotion, or behavior" (Insel 2013). Accordingly, the aim of RDoC is to create a new system of categories which refer to clusters of "genetic, imaging, physiological, and cognitive data" rather than to clusters of symptoms used to define the DSM categories.

The NIMH's position, as it was presented by Insel, provoked severe reactions, even from those who otherwise shared the concerns about the validity of the DSM categories. One of its critics, Allen Frances, posted on the website of *Psychiatric Times* the following comment: "NIMH has turned itself almost exclusively into a high power brain research institute that feels almost no responsibility for how patients are treated or mistreated in the here and now" (Frances 2013).

Eventually, the controversy "NIMH vs. DSM 5" was resolved, at least ostensibly, by a common statement issued on behalf of the American Psychiatric Association (APA) and NIMH. The statement contained the declaration that "DSM-5 and RDoC represent complementary, not competing frameworks" for achieving progress in characterizing mental disorders (Insel & Lieberman 2013). This controversy, however, was regarded by some as a sign of a crisis for the medical model of mental disorders, which both DSM and RDoC endorsed. Accordingly, the putative crisis was seen as a reason for undertaking a more radical departure from the medical model.

A year after the onset of the NIMH vs. DSM controversy, the high-profile scientific journal *Lancet* published a paper appealing for a "narrative turn in psychiatry" (Lewis 2014).[1] The author of the paper, Bradley Lewis, stated that "taking a narrative turn does not mean abandoning the biological model or any other. A narrative turn says all the models are valuable if used self-reflectively in dialogue with patients and in the service of recovery. Ultimately, it is the patient's choice which model or combination of models best fits their goals and desires" (Lewis 2014, p. 23). What is the rationale behind this appeal for a narrative turn in psychiatry? According to Lewis, all existing models that explain mental variation (including purely biological models and the biopsychosocial model as well as the psychoanalytic, cognitive, and other models) tell us different stories about mental health, using different sets of "root metaphors." It is wrong, however, he claimed, to assume that just one of these stories is true and the others are wrong. In a sense, all these stories are at the same time true and false, but they provide different perspectives on patients' treatment. Whether the suggested treatment proves to be successful or not would eventually depend on how the patient accepts the proposed story of his/her health issue.

At the time of Lewis' statements, the idea that the explanatory model of particular mental problems should be negotiated between the psychiatrist and the patient had already been popularized. It, for example, had been endorsed by Stanghellini (2013) whose suggestions for focusing the psychiatric interview on the first-person narrative of the patient will be discussed in some detail. This is because Stanghellini's views effectively demonstrate the theory-ladenness of understanding patients' first-person narratives and thus justify the suggestion that the theory-ladenness of understanding patients' narratives should be given much more attention than it has usually received.

Stanghellini' dissatisfaction with the dominance of the biological model of mental disorders,

[1] This should not be seen as a surprise. A year later the editors of *Lancet* launched an initiative to publish case reports describing "an unusual presentation of a common disease or a rare cause of a common presentation if not something completely novel" motivated by the view that stories like those "form the basis of how we learn, and how we remember" (Berman & Horton 2015, p. 1277).

which he calls the “biomedical model” or “biomedical paradigm,” is much more radical than that of Lewis insofar as Lewis is only against the dominance of this model, not against the model itself. For Stanghellini, the biological model, which assumes that the symptoms of a particular disorder form a cluster because of their common neurobiological cause, in fact works against narratization (telling a story about the formation of this cluster). This is because, according to his account, the formation of a narrative requires people to order their “actions, experiences and beliefs” by establishing meaningful connections between them, as narratives are based on meaningful connections and not merely on causal relations. Thus, the narrative turn in psychiatry for Stanghellini should consist in a radical departure from the biological model and endorsement of a model that assumes that the symptoms of a given mental disorder form “a meaningful whole” (Stanghellini 2013, p. 343).

Stanghellini also insisted that the meaningful relations between the symptoms should be extracted from the first-person narrative of the patient over the course of a dialogue (led by a non-structured interview), ensuring that the meaningful connections as they are seen by the patient are grasped correctly by the clinician. How does that work? Stanghellini tried to illuminate the process with the example of schizophrenia. A common symptom of schizophrenia shared by many patients suffering from this disorder is the delusion of external control. According to Stanghellini, however, this otherwise bizarre delusion becomes understandable if one views it as an explanation of patients’ “disturbing changes in bodily experience” (Stanghellini 2013, p. 347). Stanghellini’s general recommendation is to make sense of the patient’s seemingly absurd or meaningless beliefs and behavior by imagining the kind of world in which these beliefs and behavior would be expectable.

However, the assumption underlying Stanghellini’s “phenomenological” approach, i.e., the assumption that the symptoms of a given disorder form a coherent meaningful whole, is a strong theoretical assumption insofar as it does not rest on any direct observations or on inferences from such observations. We need this assumption, as Stanghellini himself has admitted, in order to make understandable *for us* behavior, thoughts and beliefs, which look bizarre or meaningless to “normal” people. However, if we assume, as Stanghellini does with regard to schizophrenia, that the perceptions of those who suffer from it can be seriously disturbed, why not assume that the semantic connections between perceived events can be just as impaired? What if the situation is exactly the opposite of Stanghellini’s assumption: The perceptions of those suffering from schizophrenia are intact but the way they bind them in a meaningful whole is flawed?

It’s not only the phenomenological approach to mental disorders that relies on theoretical assumptions in order to make understandable the patients’ first person narratives. It seems as if theoretical (one may call them “philosophical”) assumptions underlie all approaches to mental disorders, including the biological one, the proponents of which view it as more objective. Is Lewis then right to insist that none of the stories told by the different approaches is better than its rivals in terms of truthfulness and that the only criterion which one should use for choosing a theoretical perspective on mental health should be its capacity to produce change in the patient’s condition in a given case?

In search for an answer to this question, I’ll discuss in the next section the peculiar story of the recognition and the following attempts to define, explain, and treat the so-called factitious disorders (FD). This case study will reveal how the different attempts to make understandable the feigned narratives of the patients with FD rest on different assumptions as well as on one common assumption – that these narratives and the behavior which they support satisfy a particular psychological need. The latter, however, is a theoretical assumption which does not rest on any direct evidence. On the contrary, this assumption directly confronts the first-person narratives of many patients with FD. At the same time, patients' stories cannot be taken as a basis for an explanatory model either, due in part to their serious discrepancy with reality, but mainly because they serve purposes that hinder the recovery process. The suggestion I will try to defend is that the behavior of patients suffering from FD should not be interpreted as aimed at satisfying certain unclear (or disputed) psychological needs, but should be seen instead as self-supporting, or as an end in itself. This model of explanation, which does not directly confront the understanding demonstrated by the patients themselves, may be a more successful basis for discussing strategies to deal with the pathological symptoms, because it recommends that clinicians focus on the available pathways toward helping patients modify their behavior, rather than reinforcing the belief that problematic behaviors meet unclear and contested needs.

2. A short reconstruction of the story of factitious disorders

In ICD-10 factitious disorders (FD) are defined as "intentional producing or feigning of symptoms or disabilities, either physical or psychological" in the absence of external incentives and rewards and of "confirmed physical or mental disorder" (WHO 1992, F68.1). FD are described in a similar way in DSM-5 (APA 2013, p. 324), although there they are presented as a subcategory of "somatic symptom and related disorders" rather than of"disorders of adult personality and behavior" as in ICD-10. Notice that although FD are associated with *Intentional* lying, the goals or causes behind the intentional process are not specified. They are only negatively defined by excluding from their class any external gains or rewards. This omission, as we will see, is not accidental, and it is not motivated solely by the ideology of DSM and ISD to offer definitions based on observable symptoms only. This omission is also due to disagreements about the alleged goals and causes behind the simulation or fabrication of symptoms typical of patients with FD. Some of these disagreements, as we'll see, stem from deeper theoretical divergences.

The phenomenon of intentional lying without obvious incentives has probably been acknowledged for ages but it was first referred to by the term "factitious disorder" by Hector Gasin in 1838 (Kanaan & Wessely 2010). Since then until the middle of the twentieth century, this term had been only sporadically used. However, the situation changed in 1951, when Richard Asher gave the first more elaborated description of the phenomenon and called it "Munchausen's syndrome." Factitious disorders officially entered the third edition of APA's classification of mental disorders – DSM-III (APA 1980) – but one of their subspecies, the factitious disorders imposed on another (formerly called Munchausen syndrome by proxy), was not officially recognized until 2013 when the fifth edition of DSM was published.

Since their official recognition, factitious disorders have been a subject of persistent controversies. To understand them we should first shed light on some aspects of the context in which these disorders were recognized as such by psychiatrists.

In his 1951 paper, Richard Asher described the typical patient displaying Munchausen's syndrome in the following way: He "is admitted to hospital with apparent acute illness supported by a plausible and dramatic history. Usually his story is largely made up of falsehoods; he is found to have attended, and deceived, an astonishing number of other hospitals; and he nearly always discharges himself against advice, often quarrelling violently with both doctors and nurses" (Asher 1951, p. 339).

This general account of the typical manifestations of Munchausen's syndrome is supported by three more detailed descriptions of real cases. Asher did not propose a definite procedure for diagnosing the syndrome, he only offered some "useful pointers" that might help to recognize the condition. In addition, he made an important remark, which must be kept in mind, namely, that the fictitious stories of the Munchausen patients always contain some truth: "it must be recognized that these patients are often quite ill, although their illness is shrouded by duplicity and distortion" (Asher 1951, p. 339).

Asher ended his paper with two conclusions. The first is about the practical importance of finding a way to recognize Munchausen patients as they "waste an enormous amount of time and trouble in hospitals" (Asher 1951, p. 341). The second conclusion is about the need to find an explanation for this condition "which might lead to a cure of *the* *psychological kink*[2] which produces the disease" (Asher 1951, p. 341).

In what follows, the attempts to solve the second tasks will be traced in order to make explicit the diverse theoretical assumptions behind these attempts to find the unknown "psychological kink" and the ways these assumptions interfered with the interpretations of fictitious first-person narratives produced by patients with FD. It should be mentioned meanwhile that since the introduction of FD in DSM (first in DSM-III), the term "Munchausen's syndrome" has been reserved for a specific subtype of FD characterized by a severe and chronic course of FD associated with feigning or fabrication of mostly physical symptoms. The discussion from now on will be focused on the attempts to explain the disorders belonging to the more general category of FD as it is described in the modern editions of DSM and ICD.

In a review chapter published in 2001, Feldman, Hamilton, and Deemer reported that at least four hypotheses explaining FD had been raised before the end of the last century. The behavioral explanation of FD viewed these disorders as a result of "past social learning reinforced by current positive or negative stimuli" (Feldman, Hamilton & Deemer 2001, p. 149). The cognitive explanations referred to flaws in cognitive processing. The psychodynamic explanations, which focused on early traumatic experiences, weakened ego, and superego deficits, generally saw FD as a maladaptive intrapsychic defense. Feldman, Hamilton, and Deemer themselves argued for the fourth explanation of FD based on what they called "the self-enhancement model," which builds on the hypothesis that FD enhance or protect self-esteem.

As Feldman, Hamilton, and Deemer (2001) themselves acknowledge, although each of the available explanations broadens our understanding of FD to some extent, none of them offers a sufficiently complete understanding of the processes that lead FD sufferers to simulate and fabricate symptoms of disease without visible benefit from this. Schwartz et al. (1994), for example, have shown that social learning is, at best, only an insignificant part of these processes, because in many cases, when the necessary conditions for social learning of the sick role are present, those placed in these conditions do not assume this role and thus do not develop FD. Moreover, it is often observed that the alleged "attention seeking behavior" persists despite doctors trying to not reinforce it (Mousailidis 2019). Certain deficits in cognitive processing explain relatively well the cases of hypochondria, in which distorted perceptions of normally functioning bodily processes lead to the formation of beliefs about an underlying disease. However, in patients with FD, there is a deliberate simulation or fabrication of symptoms rather than a misperception of normal functions as pathological. Psychoanalytic explanations, although different from explanations referring to social learning, encounter similar problems: Traumatic events or circumstances that are thought to trigger FD do not lead to FD in the majority of people.

It is important to stress here that the different explanations of FD not only lead to different understandings of the processes associated with these disorders, but also influence the way psychiatrists perceive and describe the stories of their patients. The latter was recognized as a significant obstacle to the adoption of standardized protocols for writing case reports, which made it difficult to compare and consolidate the available case reports in systematic and narrative reviews. However, attempts at such reviews have been made as we shall see soon.

In 2004, Marc Feldman, one of the authors of the review chapter (Feldman, Hamilton & Deemer 2001), reduced "the ultimate puzzle" of FD to the question of "the psychological benefits of playing 'the sick roles'" (Feldman 2004, p. xvii). He also added that it is "important to understand why so little progress has been made toward the scientific understanding of this particular psychological problem" (Feldman 2004, p. xviii). Apparently, Feldman was not inclined to accept that the lack of progress might be connected to the biased formulation of the "ultimate puzzle," suggesting that the intentional simulation or fabrication of symptoms brings some psychological benefits.

What is the present understanding of FD? In one of the few systematic reviews summarizing the

[2] The emphasis (italic) is added by me, LG.

findings of 455 case reports, Yates and Feldman (2016) confess that the motivations behind FD patients are "almost always obscure," but at the same time speculate that they might include the following: desire to receive "affection and care," "adrenaline rush," "sense of control over those who are deceived." The assumption behind these tentative explanations is that intentional behavior should be driven by explicit (conscious) goals or desires transcending the immediate (proximate) ends of this behavior. But is this necessarily so? In the next section, I will point out several important advantages of considering the behavior of patients with FD as self-engaging and an end in itself, and not as determined by (still unclear) psychological "kinks" or motives, as the currently discussed explanations of this behavior suggest.

3. The advantages of understanding behavior of patients with factitious disorder as self-engaging and an end in itself rather than as directed toward the satisfaction of certain presumed psychological needs

Let's go back to Asher's 1951 article, which first described Munchausen's syndrome. Asher noted there the extreme uncooperativeness of patients with this syndrome and their fierce resistance to any attempt by clinicians to tell them that they were not in fact suffering from the symptoms that they were feigning. Almost all case reports published in the following years confirmed Asher's observations. If the patients with FD sought the sick role to attract attention, love, and care, why did they refuse the attention and care offered to them, but at the cost of acknowledging that they need them and not the treatment of the diseases or disabilities they had deliberately fabricated? A hint leading to the answer of this question can be found in the confession of a former patient diagnosed with FD, who shares the following: "I was 9 1/2 when my little sister was born. That day, my life changed forever. Not just because I was no longer the baby, but because my sister was born two months early and was diagnosed with Cerebral Palsy. She needed a lot of care and attention than the typical bay and that's when I became the invisible girl. I was emotionally abandoned by my mom and emotionally and verbally abused by my father. A couple years after her birth, I began faking things to get the attention I was missing out on. *It was something that went deeper than just getting attention. It was something deeper.*"[3]

Of course, one cannot expect a patient to account for her thoughts, feelings, and beliefs in scientific terms. As already noted, understanding of first-person narratives is always associated with interpretation or translation in terms of a known theory or theoretical framework. But do we have a theory that accounts for human behavior through some explanation other than external causes and internal psychological needs? It turns out that there are such theories.

Brian Boyd offers an original evolutionary explanation of what we can define as self-engaging behavior.[4] We often do things deliberately, not to achieve something else, not to achieve a certain effect, but for their own sake. And we keep doing these things even when they are no longer desirable simply because we can't stop or because we don't see a compelling enough alternative. The prototype of self-engaging behavior is playing. Animals also play, Boyd notes, "but *pretend* play appears to be an almost exclusively human activity" (Boyd 2009, p. 181). The explanation of pretending behavior as self-engaging is compatible with what we find documented in many case reports. Often, for instance, patients with FD are women who lead a monotonous life that does not offer any engaging activities (Feldman, Hamilton & Deemer 2001). For such women, playing the sick role turns out to be much more engaging than anything else in their daily life. Compatible with this explanation are also the descriptions of cases of spontaneous recovery of FD patients after being involved in a new engaging activity, for example, after making a commitment

[3] This is an excerpt from a patient's narrative posted by Marc Feldman on April, 2017, 2020, at Important treatment advice from a woman who recovered from Munchausen by Internet | Dr. Marc Feldman. The italic is mine – LG.

[4] Please note: A quote from Boyd's book, *On the Origins of Stories: Evolution, Cognition, and Fiction* (2009) is used as the motto of this paper

to taking care of a pet (Feldman, Hamilton & Deemer 2001).

If the hypothesis stating that playing the sick role is a form of self-engaging behavior is confirmed, this will open up new opportunities to help FD sufferers, presenting meaningful alternatives to what is currently offered within the medical model.

4. Conclusions

The narrative turn in psychiatry could not lead to success without taking into account the theory-ladenness of understanding patients' first-person narratives. This general conclusion is supported by the analysis of the so-called factitious disorders. It has been generally accepted that patients with FD strive to play the sick role in order to satisfy psychological needs. As plausible as this assumption may sound, it is not necessarily true and we do not have enough evidence to support it. On the contrary, the resistance of patients with FD to accept this assumption is a good reason to question it. This paper highlights the advantages of an alternative theoretical assumption according to which the behavior of those assuming the sick role is self-engaging and not determined by external causes or internal psychological needs. First, the idea of self-engaging behavior is consistent with the first-person stories of patients diagnosed with FD who have successfully overcome their condition. This idea also explains the typical social profile of FD patients. An important advantage is also that the prevalence of self-engaging behavior has an evolutionary explanation. And last but not least, the idea of self-engaging behavior explains the recovery of patients with FD after a successful switch to another completely different engaging activity.

References

APA (1980) *Diagnostic and Statistical Manual of Mental Disorders*, *Third Edition (DSM-III)*. Washington, DC: APA Publishing.

APA (2013) *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition*. Washington, DC: APA Publishing.

Asher R (1951) Münchhausen syndrome. *The Lancet* 1(6650): 339–341.

Berman P, Horton R (2015) Case Reports in The Lancet: a new narrative. *The Lancet* 385(9975): 1277.

Boyd B (2009) *On the Origins of Stories: Evolution, Cognition, and Fiction*. Cambridge, MA: Harvard University Press.

Feldman M D (2004) *Playing Sick? Untangling the Web of Munchausen Syndrome, Munchausen by Proxy, Malingering, and Factitious Disorder*. London: Routledge.

Feldman M D, Hamilton J C, Deemer H N (2001) Factitious disorder. In *Somatoform and Factitious Disorders*, ed. Katharine A. Phillips, 129–166. Washington, DC: APA Publishing.

Frances A (2013) NIMH vs. DSM 5: No one wins, patients lose. *Psychiatric Times*, May 10, 2013. Retrieved on Jan 1, 2021 from: NIMH vs DSM 5: No One Wins, Patients Lose | Psychiatric Times

Insel T R (2013) Transforming diagnosis. Retrieved on Jan 2, 2021 from: http://www. nimh. nih. gov/about/director/2013/transforming-diagnosis. shtml.

Insel T, Lieberman J (2013) DSM-5 and RDoC: Shared interests. Retrieved on Jan 2, 2021 from: New Mental Health Screening of U (wordpress.com)

Kanaan R A, Wessely S C (2010) The origins of factitious disorder. *History of the Human Sciences* 23(2): 68–85.

Lewis B (2014) Taking a narrative turn in psychiatry. *The Lancet* 383(9911): 22–23.

Mousailidis G, Lazzari C, Bhan-Kotwal S, Papanna B, Shoka A (2019) Factitious disorder: a case report and literature review of treatment. *Progress in Neurology and Psychiatry* 23(2): 14–18.

Schwartz S M, Gramling S E, Mancini T (1994) The influence of life stress, personality, and learning history on illness behavior. *Journal of Behavior Therapy and Experimental Psychiatry* 25(2):
135–142.

Stanghellini G (2013) Philosophical resources for the psychiatric interview. In *The Oxford Handbook of Philosophy and Psychiatry*, ed. K. W. M. Fulford, Martin Davies, Richard G. T. Gipps, George Graham, John Z. Sadler, Giovanni Stanghellini, and Tim Thornton, 321–356. Oxford: Oxford University Press.

WHO (1992) *The ICD-10 Classification of Mental and Behavioral Disorders*. Geneva: WHO.

Yates G P, Feldman M D (2016) Factitious disorder: a systematic review of 455 cases in the professional literature. *General Hospital Psychiatry* 41: 20–28.

Discussion

A continuation from the previous issue .

RESPONSE TO PLAMEN MAKARIEV'S COMMENTS

Slobodan Divjak*
Honorary President of International Philosophical School Felix Romuliana, Zaječar, Serbia
e-mail: slobodan.divjak945@gmail.com

I would like to express my gratitude to Plamen Makariev for his involvement in the discussion on the topic of liberal democracy and cultural diversity, in which he also reflected on some of the theses I presented in my text "Communitarianism, Multiculturalism, and Liberalism." With his subtle analyses of the given topic, he significantly contributed to its perception from different angles. Since he is a theorist who is well versed in the matter in question, it would be my pleasure to respond to his remarks regarding my article.

The central issue Makariev's remarks refer to is, in his opinion, my failure to distinguish between descriptive and normative communitarianism. However, my position is that this distinction, in principle, cannot be established as the concepts of communitarianism and multiculturalism developed during the 1970s and 1980s insist on a connection between the system of objectively existing particular cultural traditions of specific ethnic groups on one side, and moral and legal normative orders that aim to strengthen these traditions on the other. In this sense, a distinction can only be made between radical and moderate variations of these movements. But both include an inherent request for formal recognition of components of their cultural traditions – a request for their specific cultural and political rights to be legally recognized so that individuals can sustain their status as members of those cultural communities. Without insisting on that connection in which the specific customs and fundamental values of an ethnic group are the basis of its moral and legal order, there can be no modern multiculturalist and communitarian movements. For example, in America, which is currently a typical liberal state not based on an ethnic principle, ethnic groups exist at the societal level but are not formally and legally recognized as such. As long as these groups do not demand specific cultural and political rights but accept the current American constitutional order, they cannot be treated as part of a modern communitarian movement, regardless of how much they adhere to their customs in private life.

---

[*] About the author: Slobodan Divjak was nominated two times for membership in the Serbian Academy of Sciences and Arts. He has written eight philosophical books. His book Problem Identiteta [The Identity in Question] is the most quoted Serbian philosophical book since the Second World War. He is the founder and an honorary president of The International Philosophical School Felix Romuliana (Serbia, Zaječar).

The theoretical exponents of both communitarianism and multiculturalism are either moderate or radical critics of purely formal liberal principles because they treat these principles, together with a hypothetical social contract, as mere intellectual concepts that, while being applied, commit violence against real social life and real history. Therefore, the representatives of both communitarianism and multiculturalism are supporters of a normative reductionism with its perspective that norms are an expression of existing internal practices of narrower or wider communities, including state communities. In such a conceptual framework, norms always have an inferior character – they can only reinforce what already exists as a way of life in each community. For example, Carl Schmitt, like all leading ideologues of Nazism, insisted primarily on racial homogeneity, which either exists or does not exist; if it does not exist, it cannot be produced by any norm. Norms, according to him, can only maintain and strengthen racial purity. But even the most successful norms of this type must always be reexamined from a pre-legal point of view. Pre-normative, pre-constitutional, pre-legal social conditions are the essence of every norm.

Racism, which Makariev treats as normative communitarianism, cannot be viewed as neutral concerning the primordialist approach to a particular culture (e.g., the hypostasis of the significance of blood and soil) also found within radical multiculturalism. The Nazi ideal of the "Aryan" race as superior to other races has its origin in an ethnolinguistic notion. Modern advocates of racism generally no longer insist on the superiority of their race but instead profess a belief in racial equality. But they do inherit from more standard accounts of biological racism the thesis of interracial mixing being unacceptable in the name of preserving racial "purity." For example, a part of the French radical right does not label Algerians living in France as its citizens as racially inferior, but rather unwelcome. The message sent to Algerians is that their place is in Algeria, not France. Based on the premise that interracial mixing is unacceptable, the ideology of apartheid of a new type emerged; new because it is not propagated by members of the white race, but rather non-white. Thus, in America, African American men and women within the radical multiculturalist movement emphasize the demand for racial separation in schools.

I, of course, make a distinction not only between different types of political cultures (aside from liberal and non-liberal) but also between moderate and radical communitarianism and moderate and radical multiculturalism. Nevertheless, I think that the way of establishing and justifying norms in both orientations is the same, but with different implications for liberal democracy. Both establish norms in different specific socio-historical contexts. Different types of culture, which in a broader sense can be interpreted as a way of human life, give rise to different types of norms. Thus MacIntyre, who is certainly neither a racist nor a Nazi, points out in connection with how norms are justified: "Philosophical theories give organized expression to concepts and theories already embodied in forms of practice and types of community" (MacIntyre, 1998, 390).

I disagree with Makariev's position that multiculturalism as such cannot be sharply opposed to liberalism. Because multiculturalism is also expressed in its radical variations. This opinion is shared in a way by Will Kymlicka. Considering that polyethnic and representative minority rights are compatible with liberal democracy, he warns that the demands of minority ethnocultural groups for self-government may jeopardize the integrative function of citizenship and unified state-legal identity:

> Both representation rights for disadvantaged groups and polyethnic rights for immigrant groups take the larger political community for granted and seek greater inclusion in it. Demands for self-government, however, reflect a desire to weaken the bonds with the larger political community, and indeed question its very authority and permanence…. In the case of self-government rights, the larger political community has a more conditional existence. National minorities claim that they are distinct 'peoples', with inherent rights of self-government. While they are currently part of a larger country, this is not a renunciation of their original rights of self-government. Rather, it is a matter of transferring some aspects of their powers of self-government to the larger polity, on the condition that other powers remain in their hands. In this sense, the authority of the larger political community is derivative…. Self-government rights, therefore, are the most complete case of differentiated citizenship, since they divide the people into separate 'peoples', each with its own historical rights,

> territories, and powers of self-government; and each, therefore with its own political community as primary, and the value and authority of the larger federation as derivative (Kymlicka, 1995, 181–182).

Thus, the right to self-government, or what I call the right to political autonomy on an ethnocultural basis, may encourage minority groups to assert demands for a transfer of more state union competencies to themselves, leading to their further isolation in self-sufficient entities – in "states within states," based on the people not in the political but in the ethnic sense of the word, whose members are guided by purely ethnic rather than political criteria when voting. In these situations, the request for secession would be a natural last step. The main internal cause of the disintegration of the former Yugoslavia is the fact that the dominant concept of the nation in all its republics was not a liberal, state-territorial but ethnocultural notion. If someone in Croatia declared himself a Catholic by faith, you could accurately conclude that he was not a Serb; and vice versa, if someone in Serbia declared himself an Orthodox, you could have known for sure that he was not a Croat, which testifies to the fact that religion as an integral component of particular culture was not separated from the nation. The dominance of the ethnocultural notion of the nation in the mentioned area is proved by the fact that in the first parliamentary elections, ethnic parties convincingly won everywhere. Even after the disintegration of Yugoslavia, sharp conflicts between different ethnicities persisted in its former republics. Bosnia and Herzegovina would have long ago disintegrated into three different ethnic entities in the form of mutually independent states if there were no strong external pressure against it. In Macedonia, Macedonians and the Albanian minority have long lived not together but side-by-side. Even though after the break-up of Yugoslavia Macedonia was led by the very moderate politician Kiro Gligorov, the Albanian minority has continually demanded an increasing degree of political autonomy for itself. In Croatia, despite the mass ethnic cleansing of Serbs, there is still a great deal of hatred between the Croatian majority and the Serbia minority. I am writing this as someone who never supported Milosevic's ethnonationalism.

Both communitarians and multiculturalists reject moral monism and therefore Kant's categorical imperative, which is otherwise compatible with the pluralism of ethical conceptions derived from experience because this imperative is a negative principle and as such, is at a higher level of abstraction than the previously mentioned ethical conceptions. It has a selective function: Only those ethical conceptions that cannot pass the test of universality (e.g. racism) are rejected.

However, even among multiculturalists and communitarians, there are radical pluralists (ontological pluralism) who advocate the thesis of the incommensurability of different forms of life as a system of customs and substantial values. Thus, different forms of life or different cultures are treated as pure particularities that cannot be assessed by any universal standard, even the basic one. Such an approach can be described as value relativism because it equates all forms of communities in terms of values, even those in which people do not have equal rights in the formal sense. This approach leads its followers to the justification of internal customs and substantial values, i.e. particular norms embodied in existing practices and interactions. Of course, they are therefore forced to reject any form of transcendent, trans-contextual critique and to accept only immanent, intra-contextual critique. If any universal standard is rejected, then the question arises as to how to criticize Nazism, slavery, sharia, communities in which there are such customs as disregard for individual rights, stoning of adulteresses, burning of widows, public flogging, if these customs are an expression of long-standing internal cultural traditions. Many multiculturalists and communitarians hold "there is an irreducible plurality of valuable goods, activities, and ways of life that are conflicting and often uncombinable and that sometimes when they come in conflict with one another, they are incommensurable… these goods can (not) be assessed, compared, and hierarchically ordered by a common measure " (Mehta, 1997, 510).

As John Gray would say, in his communitarian-postmodern phase, incommensurable values are inserted into the collective identities which are not chosen but inherited:

> In political milieus which harbor a diversity of cultural traditions and identities, such as we find in most parts of the world today, the institutional norms best suited to a modus vivendi may well not be individualist institutions of liberal civil society but rather those of political and legal pluralism, in which the fundamental units are not individuals but communities. In polities that are plural or di-

> vided, the legal recognition of different communities and their different jurisdictions may well be mandated on the Hobbesian ground that it promotes peace. Such a legal pluralism is justifiable also on the Herderian ground that it allows even people who are co-mingled in the same territories or human settlements to recognize their cultural identities in the legal order to which they are subjected. (Gray, 1995, 136).

According to that standpoint, ”all groups, including majority groups, receive the right to establish ethnic autonomy to protect legally their particular cultural traditions so that legal pluralism would be introduced to such a state: each of these groups would have a legal system that would protect its particular tradition. Such a state would be temporary modus vivendi based on a feeble consensus of these groups to co-exist in the same state, whereby the legal order of the latter would be reduced to a protocolary or symbolic one. That state would no longer be a civic state. Each of these groups would live independently of groups with different cultural traditions” (Slobodan Divjak, 2018, 96).

In the text published in the *Balkan Journal of Philosophy*, I pointed to the following articulated by Michael Dusche: “radical multiculturalism claims that cultural groups, not the individual should be the yardstick for considerations of justice, because the group offers the individual the indispensable good of being rooted in the community and since membership in a culture is not voluntary, the abolition of culture leads to the uprooting of individuals.... radical multiculturalism risks falling prey to self-defeating normative relativism” (Dusche, 2004, 238).Again, not only “normative” communitarianism (in my view, without the requests for the legalization of group rights neither communitarianism nor multiculturalism can exist) but also radical multiculturalism “stand in stark opposition” with the liberal system.

Two types of universalism

Makariev’s view of universalism is different from mine. I distinguish substantialist universalism from that of a purely formal character. I associate substantialist universalism with the metaphysically understood all-encompassing universal reason (*Vernunft*). Metaphysics ontologizes reason, i.e. treats it as the universal substantial basis of all that exists. More precisely, all existing empirical forms are shown, from the metaphysical point of view, as phenomenal manifestations of the universal substance of reason. To that extent, such a concept of reason manifests itself as an all-inclusive totality within which the necessary internal relations between its moments rule. What is not included in that totality belongs to the sphere of contingency. Thus, empirical moments, from this perspective, have a secondary role for their essence lies outside of themselves in the sphere of reason as purely rational thinking. All differences are essentially reduced to differences within an all-encompassing universal reason.

According to Hegel, reason is deprived of any outside conditioning which makes it absolutely non-conditioned and independent, based in itself. Unlike understanding (*Verstand*), which is determined by empirical data that exist outside of it as an independent positive being, reason contains everything “other” within itself (the empirical, the sensual, the particular, the final), and becomes a self-positioning, self-developing, and self-uniting power. Therefore, any empirical given is nothing but a phenomenal manifestation of reason as an all-encompassing principle. It is not an empirical history but a conceptual, teleological one: What was potentially at its beginning is at its end, with all its potentials completely realized.

As a critic of metaphysical ontologizing of reason, Kant ascribes to reason a priori status. In his *Critique of Practical Reason*, relevant to our discussion as it points to reason’s relationship with the normative order – both moral and legal – Kant treats pure practical reason as reason that is unmediated by empirical data, i.e. as an autonomous rational power that, following its internal formal logic rules, produces purely formal, non-teleological, de-ontologized norms, regardless of any cultural tradition and all empirical data. Therefore, Kant has become one of the founders of the universalism of a purely formal character that has its foundation in post-metaphysical reason. Universal norms, values, and ideals derived from metaphysical reason are substantial ones by their character. It is also clear that norms, values, and ideals derived from post-metaphysical (procedural) reason are not substantial by their character. They are unsubstantiated, non-teleological, and purely formal norms, such as freedom of speech, freedom of association, freedom of religion, etc.

The problem with the metaphysical concept is that it is monistic, i.e. tends to explain all the richness of the world starting from a single point – from universal reason, the innate universal substance that only changes phenomenal forms. According to this concept, there is essentially a unity of purely rational thinking and being (*Sein*), of what is and what should be (*Sein* and *Sollen*), of the factual and normative order. Hence, according to the metaphysical system, there is only one way of life appropriate to the truly universal human nature. Of course, both communitarians and multiculturalists reject this monistic concept of essential reason because they believe that different forms of life are based on the pluralism of particular cultures as systems of customs and substantial values that are sublimates of different cultural traditions. They are also supported by the fact that history has shown that there is no broad consensus regarding any metaphysical system. I share Kant's view that, in principle, it is impossible to reach a consensus on any substantive value. In my opinion, there is no universal substantive value. Regardless of their explicit universalist pretensions, metaphysical conceptions cannot be universal.

Although they reject any universal substantial value, both communitarians and multiculturalists support substantial values that are founded in a particular socio-historical context. Therefore, they can be characterized as supporters of contextual substantialism. However, they also deny any validity of formal universalism. According to them, a notion that reason independent of tradition, relying on itself and its internal logic, could construe a universal moral and legal normative order applicable to all contexts and times, is merely a liberal chimera. However, that is wrong. Contextual substantialism and pure formal universalism are at different conceptual levels. Customs and values related to contextual substantialism can never be universal, but only particular, specifically substantial and teleological, which means that they tend to be integral aspects of people's everyday life. In contrast, norms such as freedom of speech, association, religion, etc. can only be purely formal-universal, de-substantiated, and non-theological. Since they are negative in character, they impose obligations on people that take the form not of what they should do but what they must not do, namely: They must not infringe on the formal freedom of choice of others. Hence, there cannot be a necessary connection between purely formal norms and the way of human life. Such norms cannot predetermine the content of human life, but can only limit it rationally, and these restrictions should apply equally to all. For example, Catholic religion predetermines certain aspects of worshipers' way of life: communion, fasting, seeking forgiveness, celibacy, acknowledgment of papal infallibility, etc. From the perspective of Catholicism as a religion which, according to self-understanding, is universal, these rules can be defined as prescribed positive freedoms, because living by them is supposedly an expression of human nature. The very principle of the individual's right to choose a religion cannot in any way be defined as a positive, but only as a negative freedom from which no specific religion can necessarily follow, because each of these religions is a substantialist and teleological conception.

It is correct to state that human beings as individuals or as group members can differ from each other in their individual and particular identity because there is no single universal culture, single religion valid for all people, universal form of life, etc. However, people living in a liberal-democratic state can be said to have the same abstract legal identity defined by a set of equal rights. But there is an essential conceptual difference between these two types of identities. Equal human rights enable people to maintain the acquired identity in society at their discretion and to modify or radically change it if they wish. It is necessary to distinguish the individual as a bearer of negatively determined rights that allow him to choose his own identity within certain constraints from the individual who, under the aegis of these rights, becomes the creator of his positive self-realization, the bearer of a specific identity. It is simply not possible to equate concrete, empirical citizens who are always characterized by miscellaneous forms of identity, and the notion of a legal person or "abstract citizen", i.e. a person without any cultural, religious, customary, and other specific identity forms. These terms are abstract conceptions that must not be understood ontologically. Thus:

> As Individuals, legal persons are also normatively responsible members of a community, but in a way different from ethical persons: whether as Saul of Tarsus or Paul the Apostle, a person is a member of the legal community and a person of law; each person must obey the laws and has certain rights as such a person. Ethical values are valid with respect to the particularity of a person, legal norms in respect of his or her attribute of be-

> ing a 'person' as such: they enjoy general and binding validity. They are not constitutive of identity; rather, they constitute the 'outer' framework of 'negative liberty' that both enables and limits, in the form of a 'protective cover', the positive liberty of self-realization. Persons of law do not have to identify with these norms, but they must come to terms with them: these are the norms of reciprocal respect that persons have to grant mutually and bindingly in order to enable social cooperation with legal certainty" (Forst, 2002, 263).

It seems that Makariev does not differentiate correctly between the cosmopolitan self and the liberal unencumbered one. In his text he poses the following questions, citing Jeremy Waldron: "Could we speak of a particular identity of a person who 'may live in San Francisco and be of Irish ancestry… learns Spanish, eats Chinese, wears clothes made in Korea, listens to arias by Verdi sung by a Maori princess on Japanese equipment, follows Ukrainian politics, and practices Buddhist meditation techniques?'... Isn't the latter an idealization which is too strong for postmodern times and which should not be taken seriously or, even less so, considered to be a factor in defining minority policies?"[1] Similarly, in another place, he writes: "For example, from a certain perspective, the liberal 'unencumbered' self may look like someone without true identity, someone unable to really love and care about other persons in their capacity as living human beings, rather than abstract individuals." After that, he cites Michael Sandel: "To imagine a person incapable of constitutive attachments such as these is not conceiving an ideally free and rational agent but to imagine a person wholly without character, without moral depth. For to have character is to know that I move in a history I neither summon nor command, which carries consequences nonetheless for my choices and conduct. It draws me closer to some and more distant from others."[2] Makariev finally concludes: "From the opposite perspective, there is no such thing as a self-consistent cultural identity at all…" In my view, Makariev is closer to Sandel's standpoint.

There are other authors of the multiculturalist orientation who want to establish the necessary connection between fidelity to shared understandings within a community in which we are born and our moral depth because each of us owes our identity to our social context. So Alasdair MacIntyre accentuates in his famous book *After Virtue*: "I am someone's son or daughter, someone else's cousin or uncle, I am a citizen of this or that city, a member of this or that guild or profession; I belong to this clan, that tribe, this nation. Hence what is good for me has to be good for one who inhabits these roles. As such, I inherit from the past of my family, my city, my tribe, my nation, a variety of debts, inheritances, rightful expectations, and obligations. These constitute the given of my life, my moral starting point. This is in part what gives my life its own moral particularity" (MacIntyre, 2002, 220).

It is not easy to understand the term community used by multiculturalists and communitarians. MacIntyre mentions small groups, the city, but also a nation. Nonetheless, if we exclude communities of human beings that are not defined by ethnicity (as these groups are not relevant to my current investigation), we may focus our attention on community in the sense of an ethnic community, that is, "a particular people sharing a heritage of custom, ritual, and a way of life that is in some real or imagined sense immemorial, being referred back to a shared history and shared provenance or homeland" (Waldron, 1996, 96). Of course, every ethnic group has its cultural traditions, which are a central category not only for multiculturalism but also for communitarianism.

Both conceptions of identity are prone to criticizing the cosmopolitan view, which they often equate with the liberal self. That is wrong. The cosmopolitan self and liberal "unencumbered" self are not on the same conceptual levels. The cosmopolitan self is not culturally neutral: it consists of a mélange of different cultural fragments and, as a substantial identity, it belongs to ethical or constitutive communities. It is a specific sort of concrete identity or substantial conception of the good life. Unlike the cosmopolitan self, the liberal self is culturally neutral and purely formal and, as such, belongs to the community of law. The liberal self precedes the cosmopolitan self, i.e. the legal person is prior to the ethical self, but that does not mean that individuals make their choices about the good life in a social vacuum. This preceding is not ontological but rather normative. The identity of the legal person is an external abstract facade for the ethical person which protects it from attempts by

[1] Here Makariev cites Waldron 1995

[2] Here Makariev cites Sandel 1982

others to coercively change its identity. A liberal self is a self to which no particular conception of life corresponds. As Rainer Forst emphasizes: "To recognize a person as an equal bearer of rights is one thing; it is quite another thing to recognize this person as the person who he or she is in all his or her attributes" (Forst, 2002, 27). A different approach to this question would entail an inextricable confusion between legal-negative and ethical-positive freedoms.

The essence of the liberal system is neither anti-traditionalism nor traditionalism but the preceding individual's right to choose between these alternatives. This choice is not purely rational; it can be motivated by pre-rational, emotional, and psychological factors, but it is a result of an individual's decision. Because of that, it is a realization of an individual's personal freedom to act. The dignity of human beings implies the possibility to choose between different conceptions of life. Every tradition implies some rationality. Cultural traditions are not non-reflective in character; in other words, they are not an enclosed, unchangeable Universe. Their development requires that each is subjected to rethinking and review and thus modified to accommodate new historical circumstances so they remain alive. Gadamer says that "...the closed horizon that is supposed to enclose a culture is an abstraction. The historical movement of human life consists in the fact that it is never utterly bound to any one standpoint, and hence can never have a truly closed horizon. The horizon is, rather, something into which we move and that moves with us. Horizons change for a person who is moving. Thus, the horizon of the past, out of which all human life lives and which exists in the form of tradition, is always in motion" (Gadamer, 1979, 271-273). Later, Gadamer says: "Understanding...is always the fusion of these horizons which we imagine to exist by themselves" (Gadamer, 1979, 273).

I think that thesis about particular cultures as homogeneous entities – pure particularities – is an illusion, especially in the modern era of the Internet and new forms of traveling and communication. Cultural traditions are changeable identities, but their changes have occurred within a historically continuous framework whose stabler norms help them maintain particular characters. Thus, we could say that many contemporary cultures share similar characteristics.

If we take this into account (although I agree with the opinion that observance of the customs and substantial values of liberalized cultures can lead to the creation of personalities with moral integrity), then we should be very careful when we consider the connection between a strong loyalty to our own culture and our strong moral character. Here, we must precisely differentiate between liberal patriotism and fanaticism.

It is well known that the German philosopher of the law, Carl Schmitt defends the ethnic purity of nations in his work about parliamentary democracy, as exhibited by the following passage:

> Every actual democracy rests on the principle that not only are equals equal but unequals will not be treated equally. Democracy requires, therefore, first homogeneity and second – if the need arises – elimination or eradication of heterogeneity. To illustrate this principle it is sufficient to name two different examples of modern democracy: contemporary Turkey, with its radical expulsion of the Greeks and its reckless Turkish nationalization of the country, and the Australian commonwealth, which restricts unwanted entrants through its immigration laws and like other dominions only takes immigrants who conform to the notion of a 'right type of settler'. A democracy demonstrates its political power by knowing how to refuse or keep at bay something foreign and unequal that threatens its homogeneity(Schmitt, 1988, 9).

Schmitt's radical ethnonationalism, which also includes a demand for ethnic cleansing, is not as much an expression of him elevating the German ethnicity to a superior level as it is an expression of his belief that there are irreducible differences between particular cultures. In this respect, Schmitt is one of the spiritual fathers of postmodernism, which crucially includes the tenet that the absolutization of the difference between particular cultures leads to a celebration of an absolute homogeneity inside them.

Supporters of fanatical patriotism love their state irrespective of its character (*Deutschland uber alles* – Germany above all). On the contrary, representatives of liberal patriotism love their state only if it provides all citizens with human rights as well as the individual right to choose between individualistic and traditionalistic forms of life. As I have accentuated, liberal states in some cases can recognize particular cultural and political rights of minority groups without leading to their segregation and destroying basic principles of the civic state. Comparing these two types of patriotism, we are confronted with a seeming paradox. Ontological pluralists who celebrate the concept of the irreducible difference between particular cultures argue for the ethnic homogeneity of their state because that implies their moral depth; those who criticize the liberal state because it is not inclusive enough (as it cannot admit certain cultural differences that are antithetical to liberal principles) now require that their state should be ethno-culturally "pure," i.e. if the celebration of irreducible cultural differences replaces liberalism as a social ordering principle then ethno-cultural homogeneity, rather than liberation toleration, is necessary for achieving social harmony.

Concerning minority ethnocultural rights within a state, Ernesto Laclau is right when he writes:

> These remarks allow us to throw some light on the divergent courses of action that current struggles in defense of multiculturalism can follow. One possible way is to affirm, purely and simply, the right of the various cultural and ethnic groups to assert their differences and their separate developments. This is the route to self-apartheid, and it is sometimes accompanied by the claim that Western cultural values and institutions are the preserve of white male Europeans and Anglo-Americans and have nothing to do with the identity of other groups living in the same territory. What is advocated in this way is total segregationism, the mere opposition of one particularism to another. Now, it is true that the assertion of any particular identity involves, as one of its dimensions, the affirmation of the right to a separate existence. But it is here that the difficult questions start, because the separation – or better, the right to difference – has to be asserted within a global community – that is, within a space in which that particular group has to coexist with other groups. Now, how could that coexistence be possible without some shared universal value, without a sense of belonging to a community larger than each of the particular groups in question? (Laclau, 1995, 93).

The question of human moral depth is complex. I am inclined toward a liberal understanding of traditional forms of life, although I'm critical of some aspects of my Serbian tradition. Adherence to particular customs and substantial values can, in my view, contribute to the moral depth of a human being. However, this adherence needs to be conditioned and depend on the characteristics of the norms in question. To judge these characteristics, it is necessary to appeal to purely formal moral and legal standards. Extreme ideologies like totalitarianism, authoritarianism, despotism, racism, radical Islam, and cultural traditions that do not respect individual rights or that legalize barbaric customs like serfdom, cannot pass the negative test of universalizability. Can a person who cannot choose between different options be a free and responsible human being? One response could be that the attachment people have to their community is not enforced but voluntary, and therefore an expression of freedom and responsibility. If so, why are cultural traditions protected by the appropriate legal systems? Why would human beings, ready to die defending the liberal system and its form of liberty, lack moral integrity and dignity in their actual lives? Why would a cosmopolitan renounce patriotic feeling for his country, a country in which he has equal rights as others, and at the same time be a citizen of the world? The desire to travel the world – an odyssey – does not have to exclude love for one's homeland. After all, Odysseus returned to his homeland after a great journey, where members of his family and his friends were waiting for him. Odyssey and attachment to one's home do not have to be conflicting concepts.

Plamen Makariev thinks that my opinion is an oxymoron: "The very idea of a culturally neutral culture is an oxymoron…" But this misrepresents my view. My view is that there is an essential difference between the ethnocultural understanding of the nation (common language, common ethnic origin, shared history, fated community, etc.) and the liberal understanding of the ethno-culturally neutral nation. The basic principles of the latter (freedom of speech, association, expression, etc.) are not substantial values because they do not imply concrete values human beings should realize in their lives to live a good life. The person as rights-bearer exists at a more abstract level than the ethical person. To the person as rights-bearer particular conceptions of the good life do not apply, as notions of the good life represent a linguistically, socially, ethnically, culturally, or historically mediated subject that is outside the purview of the purely formal subject. The person as rights-bearer is a part of a moral and legal normative order that forms and circumscribes ethical identities. Indeed: "Individual rights grant the ethical self constituted in the community the latitude to develop and the formal possibility to examine critically and revise this identity. This possibility does not have any direct ethical implications in the sense of particular individualistic ideals of the good life" (Forst, 2002, 26).

Although they are formal and non-teleological, these moral and legal principles are not value-free, for negative liberty has enormous value. Their values are only different in terms of their normative content, while the content of substantial values always goes beyond that of normative ones. The latter, regardless of whether they are metaphysical or contextual, always tend to be realized in actual human life. Since every metaphysical claim is that a normative system is an expression of the essence of all human beings, it treats the contexts of its norms as universal. But as already mentioned, there is no consensus about metaphysical truths; therefore, we must treat metaphysical substantial values as only a form of particular values, regardless of any ambition to universalize.

Without this essential difference between substantial values, which are always particular, and non-substantial values, which are always universal in a formal sense, it is not possible to make a fundamental distinction between systems whose norms are rooted in this or that type of community and the liberal system in which the law is independent of the ethnocultural structure of the community or state. Habermas, in reference to the notion of the self-legislation of voluntarily associated citizens who are free and equal, has said: "This idea is not 'formal' in the sense of being 'value-free.' However, it can be fully developed in the course of constitution-making processes that are not based on the previous choice of substantive values, but rather on democratic procedures. Hence, there is a justified *presumption* that the deontological idea of self-legislation or autonomy is neutral with respect to worldviews, provided that the different interpretations of the self and the world are not fundamentalist but are compatible with the conditions of post-metaphysical thinking…" (Habermas, 1998, 406).

I, of course, differentiate between the theoretical model of the liberal system and the liberal system in its empirical forms. There is no theoretical model that is fully realized. Still, such models help us to interpret an existing system adequately. The liberal system is by no means ideal. In my view, it is only less bad than other systems. It should also be noted that the liberal form of liberty must not be imposed by force; rather, the adoption of such a system must be the result of the freely expressed will of the citizens of a country. For this reason, I am a radical critic of military interventions that are allegedly humanitarian; they are, by their nature, deeply antiliberal. However, in both a liberal theoretical model and in liberal countries like America and France, bearers of human rights are not ethnocultural or religious groups but individuals. Now, I endorse particular polyethnic and representative rights of minority groups with adequate justification – in other words, if there is a specific need that is satisfied by these rights, which thereby serve an integral function in society. While these rights may be justified in liberal society, this does not mean that they are liberal rights. Indeed, we can interpret such rights only as certain necessary deviations from the norms of liberal rights, i.e. from the logic of a pure liberal system. Ethnocultural or religious groups then transcend the liberal framework of civil society so that ethnocultural principles become a part of the constitutive principles of the civil state, in which case the state then ceases to be an ethnically- and culturally-neutral nation-state. Hence, while specific minority rights may indeed be necessary, they cannot be interpreted as anything other than antiliberal in character.

I do not think that any system can exist without at least a minimum of assimilation. The entirely unassimilated system would be entirely inclusive – a utopian society. In reality, not a single state system can resolve the question of the 'Absolute Stranger.' When broaching this question, it is key to acknowledge that our central concern is not whether it is desirable for majority groups and 'outsiders' in every civilized country to establish hospitable relationships, but whether it is possible for outsiders to acquire equal citizenship. Crucially, it should be noted that the nature of the Absolute Stranger depends on the nature of the state order. But regardless of the specific system, the Absolute Stranger, justly or unjustly, cannot be a full member of civil society.In a state that defines itself in terms of race, Absolute Strangers are all human beings who belong to racial minority groups. In a state that defines itself in terms of ethnocultural heritage, Absolute Strangers are human beings of another ethnic origin.. In theocratic states, a basic condition for acquiring civil rights is belonging to a fundamentalist religion that represents the substantial base of the legal system. In liberal states, Absolute Strangers are members of anti-liberal groups, e.g. fundamentalist Islam, since their forms of life cannot be reproduced under a liberal legal system.

Some examples of such forms of life include: the institution of slavery, mass killing of people to maintain despotic rule, absence of universal individual rights, lack of women's rights, polygamy, burning widows, discrimination against allegedly 'inferior' races, an amalgamation of church and state, etc. If members of such groups refuse to abandon their radical antiliberal tradition to obtain legal status in liberal systems, they cannot acquire citizenship as this would result in two mutually incompatible legal systems existing within the liberal state. So, for example, a person cannot both embrace French citizenship and belong to a fundamentalist religion since the forms of life prescribed by religious fundamentalism undermine the liberal French social order. French citizens who defend fundamentalism must consequently abandon its radical components in order for their forms of life to be reproduced within the framework of the French formal legal order. Turks in the Turkish state in the era of Kemal Ataturk's rule had to do the same so that Turkey could become a secular state. Both of these examples illustrate partial integration within civil states through the assimilation of non-liberal groups. In these cases, assimilation to the liberal system has been only partial because these groups never completely rejected their cultural traditions; they only reformed them in the name of liberalization.

It bears mentioning, however, that it is difficult – if not impossible – to identify any particular culture which does not evolve and reform in some measure through contact with new cultural forms. Groups that are exposed to liberal social forms often undergo a process of liberalization. They may continue to follow customs and substantial values of their cultural traditions that are compatible with liberal constitutional systems while abandoning those practices that are not. In this way, their culture is reshaped by liberalism while simultaneously the forms of life that are associated with the principles of liberalism are expanded to include these newly liberalized practices. From this it is clear that liberal orders are more inclusive than non-liberal ones, and non-liberal systems are far more closed to 'the Other' in respect of extending citizenship.

What about assimilation?

Rejection of the Other by non-liberal states does not necessarily mean the absence of assimilation. Both multiculturalists and communitarians insist on tight connections between personal and communal identity. However, this thesis requires such a radical reduction of individual identity in favor of social determinants that it leads to denial of individual freedom. Proponents of this thesis claim that personal identity is carved within the communal one and therefore members of such a community do not feel its collective practices as a constraint, but rather as a result of their voluntary consent. In other words, their identification with communal identity is not a consequence of force and manipulations by communal rulers, but a consequence of their true striving to act in accordance with a concept created by pre-rationally determined shared goals and procedures of a prolonged process of hyper-socialization. Thus, MacIntyre believes that a hyper-socialized individual who is inescapably drawn to the community's view of the good supersedes an atomistic individual who can choose only arbitrarily and therefore cannot justify its choice to other free wills. However, this belief is very problematic. If we assume that it is at all possible to realize a community or a state in which there exists unity between individual and collective interests, then we must also assume that such a community can only have a sphere of positive liberties. The sphere of negative liberties cannot be present because it must be separated from the community or state's interests; in such a community or state, individual interests must be subordinated to the community or the state. This fact is denied by neither Hegel nor Marx, both of whom undertake their own totalizing metaphysical projects. The vision of community endorsed by communitarians and other thinkers who designate an identity between individual and community interests thus raises the question: is the existence of the free will that can choose what to believe, how to act, and how to live one's own life possible in such collective entities?

A free individual is one whose actions correspond to his or her choices. In my view, we can identify a person's voluntary tie to a cultural tradition only if this tradition is self-maintaining, self-developing, and freely adopted by its adherents, and is not sanctioned by any law. This does not mean that many or even most members of a community or state in which a particular cultural tradition is in fact enshrined in laws cannot also voluntarily identify with and participate in this tradition. However, history has shown that such states and communities have had their share of dissidents, i.e. those who were ready to reject these forms of life and were unwilling to change their own in order to conform. Could we not then qualify the punishing of such dissidents for their violations of laws by which the cultural traditions of state communities have been asserted as forcible assimilation? Such antiliberal state communities thus subject its members to a condition of unequal freedom. Given that the choice of cultural traditions which one practices can be either advocated or punished at an official institutional level, it is clear that members of these states have much fewer rights than others and only a limited freedom of self-determination.

The laws of state communities that protect the unequal rights of their members cannot be culturally neutral, for human rights are always connected with forms of human life. The laws of these state communities predetermine different substantial contents of the life of their members. Hence, they have at least a potentially assimilationist character. In contrast, purely formal liberal laws do not predetermine the contents of people's lives; they only constrain them.

I am not attempting to say that there are not de facto forms of discrimination in liberal systems. However, liberal systems contain an important legal component that distinguishes them from other systems: the ability to express civil disobedience against all forms of discrimination and oppression. Of course, actions motivated by civil disobedience cannot eradicate tendencies towards some forms of discrimination, but they can certainly minimize them.

Radical multiculturalism and radical communitarianism as one of its forms aim at the reintegration of what liberalism tore apart – a cultural tradition and a legal-constitutional order. Of course, there is also moderate multiculturalism, whose goal is not to reject liberalism but to correct it. As already mentioned, I am one of the scholars who believe that if social reality in some liberal-democratic countries demands the introduction of some aspects of moderate multiculturalism then that reality should be acknowledged. But one should also not forget the fact that radical multiculturalism, which undermines liberal values, is in effect moderate multiculturalism brought to its ultimate logical consequences.

Finally, unlike Makariev, I think that communism as a totalitarian regime is not "normative communitarianism," because communism is not based on any cultural tradition but on a new, non-commodity form of production.

Editor: Vanja Dragutinovic, a US Citizen of Serbian ethnic origin, bilingual in both English and Serbian languages.

References

Divjak S (2018) Multikulturalizam i moderno pravo. In *Stanje i perspektive multikulturalizma u Srbiji i državama regiona,* a collection of works from an International conference held in organization of SANU (Serbian Academy of Sciences and Arts).

Dusche M (2004) Radical Multiculturalism versus Liberal Pluralism. *Ethical Perspectives* 11: 238-249. DOI: 10.2143/EP.11.4.519089

Forst R (2002) *Contexts of Justice: Political Philosophy beyond Liberalism and Communitarianism*. Berkeley, CA: University of California Press.

Gadamer H-G (1979) *Truth and Method.* London: Sheed and Ward.

Gray J (1995) *Enlightenment's Wake*, cited in Pratap Mehta, 1997. "Pluralism After Liberalism?" *Critical Review* 11.

Habermas J (1998) A Response to My Critics. In *Habermas on Law and Democracy, ed.* Michael Rosenfeld. Berkeley, CA: University of California Press.

Kymlicka W (1995) *Multicultural Citizenship*. New York: Oxford University Press.

Laclau E (1995) Universalism, Particularism, and the Question of Identity. *The Identity in Question* 61: 83–89.

MacIntyre A (1998) *Whose Justice? Which Rationality?* Notre Dame, IN: University of Notre Dame Press.

MacIntyre A (2002) *After Virtue.* Notre Dame, IN: University of Notre Dame Press.

Mehta P (1997) Pluralism after Liberalism? *Critical Review* 11.

Schmitt C (1988) *The Crisis of Parliamentary Democracy*. Cambridge, MA: MIT Press.

Waldron J (1996) Minority Cultures and the Cosmopolitan Alternative. In , *The Rights of Minority Cultures,* ed. Will Kymlicka. New York: Oxford University Press.

Book reviews

CONNECTING THE DOTS

A Review of
Early Analytic Philosophy and the German Philosophical Tradition
(BLOOMSBURY ACADEMIC, 2020)
by Nikolay Milkov

Kamen Dimitrov Lozev
South-West University, Philosophical Faculty
E-mail: kamenlozev@abv.bg

This latest book authored by Professor N. Milkov I find to be a *challenge*. Indeed, while reading it, and upon finishing the final page, I could not rid my mind of the word 'challenge,' which, in truth, summarizes the strong impression the book left me with. This review, therefore, is an attempt to rationalize my spontaneous impression of the book *as a challenge*.

***

I think the book challenges the reader in a number of ways, but I shall concern myself with only two of them. First, to me, the book is a challenge because of its *breadth*. What I mean is this: in the Preface, Milkov tells us the story of how his personal interest in the topic "was aroused." We see him some fifteen years ago in Pittsburgh participating in seminars on the history of Analytic philosophy, Frege, and Sellars, where, according to his own words, he was "struck by the profound influence that W. V. Quine has had on American philosophers" (p. viii). The history of Analytic philosophy had obviously been approached in the context of the notorious *split* between Continental and Analytic philosophy. From that moment on, Milkov firmly decides not simply to oppose "the divide in contemporary philosophy but to contribute to the annihilation altogether of this particular idea. This really is a *grand* ambition, and indeed a broad and difficult aim to achieve.

However, this is *not* what Milkov tackles in the book under consideration. His focus is much narrower and his effort may be better viewed as an aspiration to demonstrate the *unity* of the philosophical process in a very concrete aspect. What is the nature of this unity and how does Milkov expound it? His plan is to show the continuity of the philosophical process in its "German dimension" by inquiring into the *link* (and, obviously, arguing for it!) between Germanophone philosophy and the emergence of Analytic philosophy. It is not for nothing that one of Milkov's first /sub/chapters is entitled: *Early Analytic Philosophy as Continuation of Mainstream German Philosophy* (p. 4). The first three words of the title Milkov associates with Russell, Moore, and Wittgenstein and, roughly speaking, this is how he delineates the first (early) period of the development of Analytic philosophy, distinguishing it from its middle and late(r) periods. Now the reader will perhaps think that the last three words of the title designate the Germanophone philosophy of the nineteenth century; however, Milkov's strong ambition takes him to investigate the roots of early analytic philosophy back to Leibniz (or even earlier – back to Plato and his *peirastic dialectics*), through Kant and Hegel, and up to Frege, Husserl, and the Greifswald objectivists.

This is a *big challenge* for the reader. One has to be very well versed in the 'classical' German philosophers starting from Leibniz on – Kant, Hegel, Trendelenburg, Lotze, Frege, Husserl, and others, together with the contemporary literature on them. This, to me, poses a huge difficulty – that is, of course, *only* in case one wants to deeply understand Milkov's ideas and arguments and to keep up with his pace of narration.

The challenge is made even greater because in his analyses Milkov covers historical periods and philosophical figures and their schools but in a *specific way*: We are not faced with a narrative typical of any history of philosophy; instead, Milkov introduces *selected* episodes, movements, and thinkers whose ideas and writings serve his main goal or thesis stated in the chapter's title (*Early Analytic Philosophy as Continuation of Mainstream German Philosophy*). Milkov tracks down the leading tendencies, positions, views, ideas, techniques which Russell, Moore, and Wittgenstein held, developed, or utilized in their work and demonstrates that they have come, or have been appropriated, from Germanophone philosophy in the form of influences and debates, criticism and rejections, exchanges of letters and ideas, agreements and disagreements. The purposeful 'selectivity' of Milkov's analyses and discussions, which at places I find understandably (and I am ready to say *excusably*) brief, render the unavoidable *kaleidoscopic* character of his narration.

There is no doubt that the manner in which Milkov's exposition unfolds is really interesting and intriguing: The reader is immersed in an atmosphere of intense stimulation of the mind, and even the imagination, and he often feels himself enlightened on certain famous or well-known episodes, teachings, and ideas interpreted by Milkov in a new way.

This undoubtedly is a positive feature of the book; still, the thorny task of developing a good understanding of Milkov's text is left up to the reader. I can elucidate this by way of example. Personally, I was most intrigued and interested in Chapter 12 (G. E. Moore and Johannes Rehmke) (p. 152), dedicated to the German "home-grown version of a proto-analytic philosophy" in the face of the Greifswald school (p. 9), and especially in Dimitri Michaltchew. It took me about a week to gain sufficient knowledge of Moore's review on Michaltchew's book. It turned out that after 1989 and the drastic revision of our history (and the Humanities altogether in Bulgaria), Moore's review had been translated (1991) and much discussed within the Bulgarian philosophical community in the context of the overall influence which Michaltchew had exerted on the ('national') philosophical development/process (Dimitrova 2019). All this was really enlightening – but my week was gone.

***

Milkov's book offers another sort of challenge related to the very *substance* of some of his interpretations and arguments. Some of them are indeed perplexing in comparison to the so-called 'received view' to which our mind usually clings. Here, before giving an example, I want to stress something which impressed me most.

I have never read in anyone else's writing the *truth* about the profound influence which Leonard Nelson had on Karl Popper. In the book (Chapter 13, p. 166) Milkov has analyzed this influence with precision and depth (see also Milkov 2012). I had the rare opportunity to apprehend this myself when in 1992 I held several discussions with Sir Karl in my capacity as his Bulgarian translator of *The Open Society and Its Enemies*. Among the first books Popper showed me (mainly translations of Plato's dialogues produced over the past two hundred years) was a book of Leonard Nelson, and I well remember how inspired he was when speaking about Nelson and his own interest in Nelson's writings, and the subsequent influence Nelson exerted on his thought.

In places, Milkov's book offers interpretations and conclusions which go against some well-established and popularly shared views. I refer here especially to Hegel and his influence on Russell and Moore. I am in complete agreement with Milkov on the great influence of Lotze (I strongly emphasize this in my book on Frege (Lozev 2019)), and yet I find it perplexing to read the title of a (sub)chapter (p. 94): 'Russell Misinterprets His Own Philosophical Development.' Milkov's thrust here is to convince us that Russell, and obviously also Moore at the end of the nineteenth century, *wrongly* thought of themselves as (neo-)Hegelians. Now we all remember the famous 'liberation' episode, excitingly described in Russell's *Autobiography*, when he and Moore *liberated* themselves… from *what*? Or from *whom*? I think the *received view* says – from the German idealism, from Kant and Hegel. I am ready to concede that while reading these pages from Russell's *Autobiography* one ponders the nature of the British idealism which had originated under the influence *allegedly* of Hegel. This *nature* was so specific that one would sometimes have difficulties in recognizing Hegel in the British idealists. In my opinion, Russell and Moore were *Bradleyan* Hegelians or *Bradleyan* idealists who liberated themselves from this *British/Bradleyan* idealism/Hegelianism, though Russell unequivocally says in the *Autobiography* that he had read the "Master" (Hegel) himself and rejected him in full. Questions arise here as to whether Russell read Hegel in German (in this language, it is well known, Russell was perfectly fluent) or in the English translation (interpretation). It would have been interesting if Milkov had investigated this broader issue of how *Hegelian* the British neo-Hegelians actually were. Another question is this: If we assume that Russell, and obviously Moore, had read Hegel in the original language, and at least had an idea of his *true* or *authentic* teachings, and if Hegel's ideas really had some "contribution" to the emergence of early analytic philosophy, why were Russell and Moore so fiercely opposed to Hegel and Kant and why did they never acknowledge the positive influences of these two thinkers on the development of their "new philosophy"?

These and other issues, however, might be interesting for readers like me (or maybe *us*); but are they in line with the overall topic of the book and the author's strategy to defend his theses? I am doubtful about this; which means that Milkov has every right *not* to deal with such issues (otherwise, he might have written several volumes on the topic). What I am not doubtful about, however, is that Milkov has written a very interesting, informative, and stimulating book which only deserves our high praise.

I find the book challenging at places, and this is the best one could hope for from an interesting book. I certainly learned a lot, and after finishing the book, the lacunae in my knowledge of the history of early analytic philosophy have become clear to me. Therefore, my concluding lines are a strong recommendation of the book to all interested in the history of analytic philosophy and its continental roots.

References

Dimitrova N (2019) Balgarskoto remkeanstvo na svetovna stsena, v: Dimitrova, N. *Prostranstva na identichnost – refleksii varhu balgarskata filosofska kultura*, Sofia: IK „Sv. Ivan Rilski" (Institut za balgarska filosofska kultura). [Dimitrova N (2019) Bulgarian Rehmkeanism on the World Stage, in: Dimitrova, N. *Spaces of Identity – Reflections on the Bulgarian Philosophical Culture*, Sofia: „Sv. Ivan Rilski" Publishing house.]

Milkov N (2012) Karl Popper's Debt to Leonard Nelson. *Grazer Philosophische Studien,* 86: 137–156.

Lozev K (2019) *V navecherieto, ili drugata revolyutsia: Gotlob Frege*, Blagoevgrad: BON. [Lozev K (2019) *In the Eve, or the Other Revolution: Gottlob Frege*, Blagoevgrad: BON.]

Variations on Process Metaphysics

[Review of the volume *Variations on Process Metaphysics* (eds. Alexander Haitos and Helmut Maassen). Cambridge Scholars Publishing, 2020, ISBN 978-1-5275-4832-9]

Vesselin Petrov
Institute of Philosophy and Sociology, Bulgarian Academy of Sciences
e-mail: petrov.vesselin@gmail.com

The book collects essays on different aspects of Whitehead's philosophy that include philosophical questions concerning the notion of life, aesthetics, political science, neuroscience, comparative and historical studies, etc. The first question for us to answer is why the title of the book is "Variations…"? Variations is a musical term, and applying this term in a philosophical context allows us to widen the perspective of what could be done in the field of process philosophy. As we know, speculative philosophy is the endeavor to frame a system of general ideas in terms of which every element of our experience can be interpreted. The general scheme of ideas offered by Whitehead's speculative philosophy is capable of variations, but there is always a certain thematic anchor for the variations. Such an anchor can be found in the present volume which constitutes the relevance of Whitehead's philosophy for our time: the idea of process as a living idea. The book presents nine different variations. Let me address these variations one by one.

The first variation is "Reception of Whitehead's Process Philosophy by Neuroscientists" by Elmar W. Busch. Its aim is to answer the question of "whether his [Whitehead's] metaphysics is received and recognized in neuroscience" (1). The basis of achieving this aim is Whitehead's own statement that "the philosophical scheme should be … applicable …. Here 'applicable' means that some terms of experience are thus interpretable" (PR, 3). In this case, the author suggests applying Whitehead's metaphysical scheme to unresolved neuroscientific problems. One such unresolved problem is that of "how the many individual phenomena of the brain combine to create a unified experience" (2). The author, using many quotations from Whitehead's major works, first provides an overview of Whitehead's brain theory, which concerns life, emotions, conscious events, consciousness, the idea of a person, and neurology. But a problem is that Whitehead has not examined the specific questions systematically. Among the unanalyzed questions are: "how Whitehead's events, and especially conscious events, can be localized within the brain" and "why do the billions of simultaneous events in the brain lead to the subjective feeling of a unified experience"? (4). The author presents a quick overview of some methods in neuroscience that might be helpful for achieving the general aim of investigating these sorts of questions. The given examples illustrating such methods indicate that "neuroscience contributes thrilling results and interpretations to the discussion" (10). It's a pity, however, that prominent contemporary neuroscientists do not reference Whitehead. According to the author, this is possibly due to the fact that it is difficult to demarcate the bounds of Whitehead's own theory and appropriate these concepts within another context (10). The author also emphasizes that Whitehead's process philosophy can be characterized as proto-pan-psychism (12) and that many important representatives of analytic philosophy of the past 20 years assume the possibility of panpsychism. The chapter ends with a suggestion about future developments in the philosophy of neuroscience and underlines "the necessity of metaphysical theorizing" (13).

The second variation is "On the Edges of *Heimat*: A Translation of Process Philosophy into Political Science" by Joachim Klose. The reason for choosing this as the topic of the chapter is the author's conviction that people have made a mistake in thinking that the market can regulate everything. The author thinks that it is necessary to have a "balance between mobility and being open to the world on the one hand, and being tied to a locality on the other" (15). For that purpose, he uses the German term *Heimat* – a word meaning homeland, with the connotation of cultural roots. He considers *Heimat* as an example of Whitehead's actual occasions. So, the aim of the chapter is to "uncover the connections between the discussion of *Heimat* and Whitehead's process philosophy" (16) and, on that basis, to give a proper definition of *Heimat* by which present societal developments pertaining to globalization can be interpreted. *Heimat* "is not a region … but rather our internally conceived relationship to the world" (22). The author observes that *Heimat* is "the sum of the positive prehensions of individual localization," and that in general it denotes "the socio-cultural foundation of individual conduct" (23). He argues that the developmental process of *Heimat* correlates to the development of Whitehead's actual occasions (24). "Physical and mental poles are related to each other through the societal bond of the community, meaning the aspects of life. This is the third dimension of *Heimat*." (34). "It is not a static feature of locations, but it denotes the development of events … which can be described analogously to Whitehead's actual occasion" (44). "With the Heimat … a conceptual framework is now available which places the focus on the dynamic relations in the developmental process of societies and makes them understandable" (ibid).

The third variation is "Rethinking the Concept of Organism and Freedom as the Foundation of the Process of Life: A Comparison between A.N. Whitehead and H. Jonas" by Regine Kather. Since the beginning of the twentieth century, many philosophers and scientists have focused on the concept of life in attempting to overcome the conception of a dualism of matter and mind introduced by Descartes. Among them is Whitehead – with his concept of organism that is applied to both inorganic and organic entities. Another philosopher who has shown interest in analogous problems is the German Hans Jonas who was familiar with Whitehead's philosophy of nature, but he criticizes Whitehead for inadequately characterizing the phenomenon of life. According to Jonas, Whitehead "is unable to take into account the existential and ethical meaning of life and death for any animate being" (50). In the present chapter, the concept of organism is taken as a background that both authors – Whitehead and Jonas – have shared as a guide for the analysis of life. The aim of the chapter is to compare both authors and to investigate "what they have in common and how far they differ" (51). The center of both philosophers' views is a revised concept of organism. For both of them animate beings cannot be conceived of as substances independently of relations to the world: An organism is essentially constituted by its relation to otherness. According to Jonas – in contrast to Whitehead – "the acts of self-transcendence that tie an organism to its environment imply at least rudimentary forms of judgment" (52). On the other hand, for both Whitehead and Jonas an organism is not just embedded in relations that are pre-determined; on the contrary, it musto generate the relations by its own activity. Indeed, "… whereas Whitehead focuses on the self-constitution of a unique organism, Jonas again concentrates on the self-centeredness of the feeling and knowing self" (55). Similarly to Whitehead, Jonas argues for the co-evolution of consciousness and freedom. For both thinkers the process of evolution proceeds discontinuously. However, Whitehead overcomes the dualism of body and mind for inorganic events; Jonas overcomes it only for living entities. The same holds true for the concept of freedom: Whitehead extends the range of the concept of freedom to inorganic processes, while Jonas argues that freedom implies at least a trace of inwardness. Jonas' concentration on living creations leads to another difference between both authors: "[F]or Jonas only living entities and the biosphere have an intrinsic value, whereas inorganic processes … have only a tendency to inwardness" (65). "Whitehead argues that the whole range of nature, inorganic processes included, has an intrinsic value that is independent of human interests" (ibid). According to Whitehead, the struggle for survival cannot exclusively be based on the fight against others; it is also grounded in manifold forms of cooperation. In contrast to the ethics of Whitehead, Jonas tends toward a biocentrism focused on living creatures and their environment. In addition, Whitehead establishes the basis for an aesthetics of nature, which is completely absent in the philosophy of Jonas (66). In the conclusion of the chapter, the author argues that "the philosophy of nature of Whitehead and that of Jonas may therefore be seen as complementary perspectives" (68).

The next variation is "Leibniz and Whitehead on a State of Change and the Constituting of Motion" by Tamar Levanon. This variation focuses "on the common ground between G.W. Leibniz and A.N. Whitehead in their analysis of the phenomenon of motion" (72). It highlights the striking similarity between the two thinkers and emphasizes "the crucial role played by the idea of a 'state of change'" in both of them (ibid). The author first considers the problem of the phenomenon of change: whether change is real or only apparent, and what role is played by this concept of a "state of change." The views of a number of famous philosophers are considered, including: Parmenides, Aristotle, Russell. The author focuses next on Leibniz's and Whitehead's analyses of motion as transition or a state of change. Specifically, the question that is investigated is: "Is it possible to provide a non-reductionist account of motion, one that avoids an explanation in terms of mere disposition in space?" (76). A comparison between Leibniz and Whitehead is made in answering that question, and their similar understanding of the role that metaphysical presuppositions play within scientific inquiry is stressed, especially their views that "metaphysics and science are simply two different perspectives on reality" (77). For both Leibniz and Whitehead, "metaphysics and science are interwoven" (79); and "action is fundamental to being or else change would be meaningless" (ibid). Another similarity between them is that they both maintain a relational picture of space and time; for them motion is also relative. They "incorporate the idea of a state of change within their analyses of motion" (81). Leibniz, who argues that motion consists only of a relation, "preserves the idea of absolute motion alongside that of relational geometrical motion" (ibid). On the other hand, Whitehead denies that the reality of motion can be defended by the idea of absolute motion. One of the differences between Leibniz and Whitehead is that "while Leibniz assumes that there is motion at an instant and that bodies are actually never at rest," for Whitehead "the primordial elements in nature cannot be constituted, or analyzed as instantaneous" (84). Another difference between the two thinkers is "the contrast between Leibniz's monads, which can change … and Whitehead's actual entities, which cannot" (84). In a word, the idea of a state of change is at the heart of Leibniz's and Whitehead's analyses of motion. There is more than one way to formulate the idea of motion as a state of change and the comparison made between Leibniz and Whitehead is an example of that.

The fifth variation in the book is "Samuel Alexander and Alfred North Whitehead: a Comparison" by Vesselin Petrov. The aim of this variation is to investigate Samuel Alexander's position in the dilemma of continuity-discreteness and to make a short comparison of his views on that topic with those of Whitehead. The chapter begins with a short description of the various views in contemporary philosophy of Samuel Alexander. This review leads to the conclusion that it is necessary to continue exploring Samuel Alexander's philosophical views and their role in contemporary philosophy. One such investigation begins with a review of the main features of Alexander's metaphysics. It is emphasized that the basis of his metaphysics is his conception of space-time and that he is a kind of non-substantivalist – "a fact that usually characterizes process philosophers" (96). Alexander's position is then considered in terms of the dilemma of continuity-discreteness, and it is argued that "he defends the side of continuity" (97). He describes the difference between the view of continuity as infinite divisibility and of continuity as connectedness as the difference between compactness and density; it is clear that the second of these descriptions of continuity is deeper than the first one (as infinite divisibility). Next, a comparison between Alexander's and Bergson's views on space and time is made. Finally, a comparison between Alexander's and Whitehead's views on space, time, and continuity-discreteness is made. It is argued that concerning the dilemma of continuity-discreteness, Alexander "is more on the side of Bergson" (102), though he criticizes him for the "spatialization" of time. The general conclusion of the investigation in this chapter is that "there are two basic and alternative approaches to process thinking. Samuel Alexander belongs to one of these approaches, being to some extent closer to Bergson, while Whitehead belongs to the other" (108).

The sixth variation is "The Aesthetic Truth" by Helmut Maassen. It is devoted to the place and role of aesthetics in process philosophy, and especially in Whitehead's philosophy. The investigation begins with its role in process metaphysics. First, the author enumerates process philosophical publications on the role of aesthetics in Whitehead's philosophy to reach the conclusion that the previous publications do not explain this role deeply and thoroughly enough. One of the main shortcomings in these publications is the restriction of aesthetics to art. Of course, there are also publications which argue that aesthetics plays a central role in Whitehead's metaphysics: "[A]esthetics is regarded as a basic concept in Whitehead's cosmology" (113). The present chapter stresses that for Whitehead, aesthetics is not just a concept related to fine art but has an ontological status(116). The investigation continues with the role of aesthetics in religion and ethics. "The central role of Aesthetics in Whitehead's Metaphysics characterizes the aesthetic turn in the interpretation of Whitehead's metaphysics, not only that of his aesthetics" (122). The chapter ends by emphasizing the fact that there are many references to aesthetics even in Whitehead's lecture notes, *Philosophical Presuppositions of Science* (1924). Perhaps the only philosopher besides Whitehead to stress the central role of aesthetics in metaphysics was Charles Sanders Peirce (126).

The seventh variation is "Aesthetics of Existence: A Whiteheadian Existential Sensibility" by Alexander Haitos. In a sense, it is a continuation of the investigation from the previous variation. The aim of the text is to emphasize the aesthetic dimension of experience which is an important factor in Whitehead's metaphysical description of the world. In his works, Whitehead stresses that the vague aspects of our experience have great potential for philosophical discovery. This is true also of the vague feelings within and of our bodies. Whitehead has an existential sensibility, meaning his work was imbued with a sensitivity towards the inner life of human beings, the values and meanings we experience. He finds our situation in the world to be characterized predominantly by aesthetic values. Thus, the purpose of the chapter is to describe how our experience is characterized by these aesthetic values. This project raises certain questions, including: How might Whitehead's systematic way of thinking shed light on our existential situation? And how might this help us to think about and respond to the precipitating crises of modernity? In responding to these questions, the author first investigates Whitehead's writing on religion, focusing specifically on Whitehead's discussions of religious experience and its metaphysical implications. Next, he turns directly to the broader idea of aesthetic experience. All order finds its root in the inheritance of a world and its integration towards the self-creation of a new entity – a enjoyed value that feels itself to be of the world and is enjoyed. That is why Whitehead writes that all order is aesthetic order. This also includes moral order, which merely consists of certain aspects of aesthetic order. A dominating environmental order is an aesthetics of existence. For Whitehead, our existential situation is fundamentally aesthetic. A creature's response to its world – how it understands, moves, behaves, etc. – is aesthetic.

The next variation is "The Relevance of Whitehead's Category of Actual Entity to Ecological Civilization" by Maria-Teresa Teixeira. The author first stresses the importance of the category of 'actual entity' and reiterates its meaning. Actual entities are "the immediate result of creativity; they are units with a novel and unifying character that emerge from the many antecedent occasions" (149). These unitary elements are atomic "because they lack divisibility, and because they form themselves as particular, individual units" (149-150). Actual occasions are also "complex, individualized, and changeless" (151). They create continuity as they become, but becoming is not continuous. The author observes that many process philosophers have insisted on the importance of oneness and unity; in fact, this is in accordance with Whitehead's notion of actual entity and its relevance to the development of harmonious living on our planet (154). In this regard, she quotes Sri Aurobindo and emphasizes that "ecological civilization could greatly profit from learning about the creative and holistic dynamics of actual entities" (155). Many examples are given demonstrating that "the category of actual entity can shed some light on the necessity to reconsider Nature as a unified and complex whole" (158).

The last variation in the volume is "The Notion of Balance: Overcoming Alienation from Nature" by Denys Zhadiaiev. It also deals with ecological problems and environmental issues focusing on the problem of alienation from nature. However, the general task of the text is a metaphysical one. Its goal is "to find a common ground where human nature can flourish with itself and other organic forms of existence" and the objective is "to analyze and describe the notion of balance introduced in Whitehead's philosophy of organism" (161, ibid). The author first compares the Bergsonian framework with Whitehead's one. He suggests considering the notions of order and satisfaction as closely related to the notions of control, procession, understanding, vision, rationalization, and definiteness. He begins the investigation with the notions of order and satisfaction, then focuses on investigating balance as the final phase of a process and a new process of concrescence. According to him, "balance is both the final phase (satisfaction), in Whiteheadian terms, and the beginning of a new process of concrescence. … it has no static character" (170). In addition: "Balance is neither a characteristic, nor a substance. It is rather something pre-established" (ibid). Next, the author gives some illustrations of balance. First, he generalizes, stating that balance is something pre-established; second, that the focus on only one limited part of reality is fraught with the consequences from other parts; third, that human beings usually try to avoid or escape the outcome of balance (174). The conclusion of the chapter holds that "[t]he definition and illustration of the idea of balance can help create proper conditions for awareness and an attitude toward nature, not merely as to another being, but as to the entirety of human existence, where both nature and humanity are reflections of each other" (175).

Resuming what has been said about the variations of the volume, it can be remarked that it could be better to have variations that are much more closely connected to each other, but nevertheless even in their present form they give the reader a deeper understanding of Whitehead's organic philosophy and of its relevance to our present day problems.