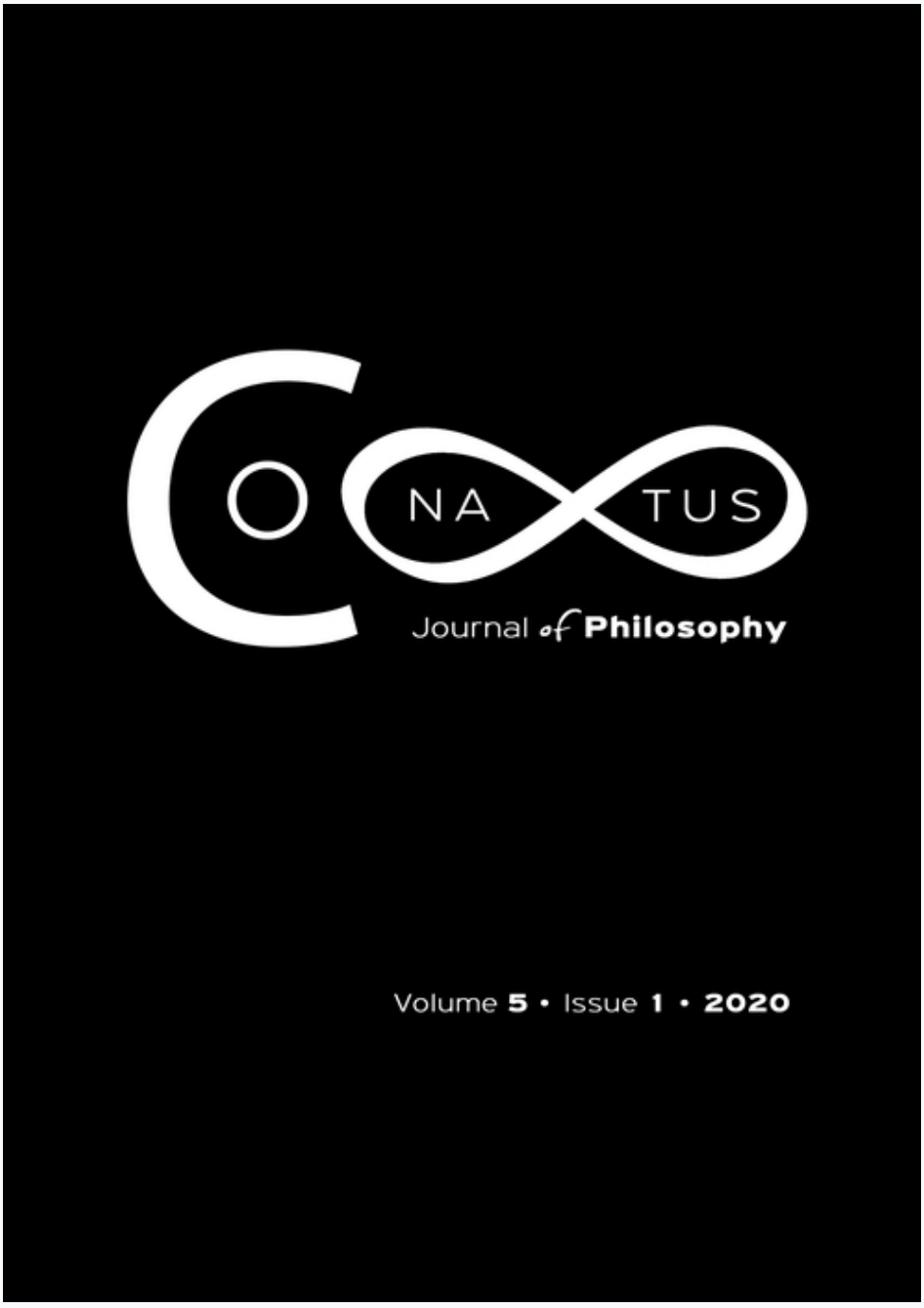
Conatus - Journal of Philosophy

Vol 5, No 1 (2020)

Conatus - Journal of Philosophy



Journal of Philosophy

Volume 5 • Issue 1 • 2020

Volume **5,** Issue **1** • **2020**

ρ-ISSN: 2653-9373
e-ISSN: 2459-3842

p-ISSN: 2653-9373
e-ISSN: 2459-3842

Contact information
School of Philosophy
7th floor, Office 746
University Campus, 15703 Zografos, Athens, Hellas
e-mail: conatus@philosophy.uoa.gr
http://conatus.philosophy.uoa.gr
https://ejournals.epublishing.ekt.gr/index.php/Conatus

contents

Articles

NEFARIOUS PRESENTISM: A RECOURSE TO PRIMITIVISM **9**
Aderonke Ajiboro

THE THEFT: AN ANALYSIS OF MORAL AGENCY **27**
Gerard Elfstrom

WHAT MAKES FREE WILL FREE: THE IMPOSSIBILITY OF PREDICTING GENUINE CREATIVITY **55**
Nikos Erinakis

IDENTITY THEFT: A THOUGHT EXPERIMENT ON THE FRAGILITY OF IDENTITY **71**
David Menčik

MAKING SENSE OF DIGNITY: A STARTING POINT **85**
Filimon Peonidis

BALANCING CONSCIENCE: A RESPONSE TO FERNANDES & ECRET **101**
Ian Wolfe & Maryam Guiahi

Interview

ANTIQUITY REVISITED: A DISCUSSION WITH ANTHONY ARTHUR LONG **111**
Anthony A. Long & Despina Vertzagia

articles

A. Ajiboro · *Conatus* 5, no. 1 (2020): 9-26
DOI: http://dx.doi.org/10.12681/cjp.21494

Nefarious Presentism: A Recourse to Primitivism

Aderonke Ajiboro
Obafemi Awolowo University, Nigeria
E-mail address: aogundiwin@gmail.com
ORCID ID: https://orcid.org/0000-0002-5622-1158

Abstract

Presentism is one of the various views in the discourse on the existence of time and spatio-temporal reality which holds that only the present is real and also that only present things exist. Neil McKinnon characterizes presentism in some ways that are all problematic, although he claims that the most appealing of all is the statement that "only present entities exist." This view permeates all thoughts about presentism, and it has led to problems about the formulation of presentism. The link between accepting the existence of a temporal part (present) and the events that have that part as its spatio-temporal reference creates a hub of debate among presentists, and this raises a lot of issues not just in metaphysics, but in other areas of philosophical discourse as well. Tallant and Ingram take a challenging position on this issue as presentists in their own right. For them, the requisite status of a presentist properly so-called should be of a commitment to the reality of the present exclusively. In this paper, I engage the views of Tallant and Ingram on the problems of presentism such as triviality and truthmaking as regards ontological implications. I will argue that the avoidance of ontological commitment in nefarious presentism does appear to avoid the problem of truthmaking, which implies avoiding an analysis of truth in order to solve the problem of truthmaking. I will also argue that this avoidance to address the principle of un-analyzability of thisness is a recourse to primitivism.

Key-words: *time; analyzability; nefarious; upstanding; presentism; existence; thisness*

I. Introduction

The metaphysical representation of reality is one of the problems in discussions on the philosophy of time. Presentism[1] attempts to address this problem, and this brings forward a typical challenge. This is succinctly put by Simon Keller's view of presentism as,

[1] Thomas M. Crisp, "On Presentism and Triviality," in *Oxford Studies in Metaphysics*, vol. 1, 15-20, ed. Dean W. Zimmerman (Oxford: Oxford University Press, 2004), 16.

> [...] the belief that only present things exist. If something doesn't exist now, says the presentist, then it doesn't exist at all.[2]

There seems to be an assumption that things do exist in the present. However, such commitment does not sufficiently explain the ontological persistence or otherwise of things when the present is not present or no longer present, as regards the truth of past events or things. This commitment can be said to be the basis of the distinction that Tallant and Ingram made of the two factions of presentists.[3]

> Upstanding presentists aim to meet the challenge, posting presently existing truthmakers for truths about the past; nefarious presentists aim to shirk their responsibilities, using the language of truthmaker theory but without paying any ontological price.[4]

The interaction of time and reality is at the center of the debate on presentism. Due to there being an assumed ontological implication that presentism points at, not of the present but of events or actions that can be said to have happened in the past but referred to the present. The link between accepting the existence of a temporal part (present) and the events that have that part as their spatio-temporal reference creates a hub of debate among presentists[5] and this raises a lot of issues not just in metaphysics but in other areas of philosophical discourse. Tallant and Ingram take a challenging position on this issue as presentists in their own right. For them, the requisite status of a presentist properly so-called should be of commitment to the reality of the present exclusively, rather than a commitment to establish the ontology of events in other spatio-temporal parts, such as the past and the future.[6]

II. Nefarious Presentism

The basis for the debate between the upstanding presentists and the nefarious presentists is about the establishing or justification of temporal ontology

---

[2] Simon Keller, "Presentism and Truthmaking," in *Oxford Studies in Metaphysics*, vol. 1, 83-104, ed. Dean W. Zimmerman (Oxford: Oxford University Press, 2004), 84.

[3] Jonathan Tallant, and David Ingram, "Nefarious Presentism," *The Philosophical Quarterly* 65, no. 260 (2015): 355.

[4] Ibid., 355.

[5] Neil McKinnon, "Characterising Presentism," in *New Papers on the Present: Focus on Presentism*, 13-30, eds. Roberto Ciuni, Kristie Miller, and Giuliano Torrengo (München: Philosophia, 2013): 14.

[6] Tallant, and Ingram, "Nefarious Presentism," 356.

in talking about spatio-temporal events in the presentists' language.[7] Upstanding presentism is the view that there is a means of providing the ontological ground required for the truth of propositions about the past.[8] This view intends to solve the problem of temporal ontology in the presentists' language by attempting to provide truthmakers regarding past events.[9]

On the other hand, nefarious presentism holds that the presentists' language can strictly quantify over present things and also that events which are no more in the present or not yet in the present are translatable. Such present tense quantification, with no commitment whatsoever to the actual existence of those events that are not available, satisfies the requirement of temporal ontology for a translation.[10] In other words, the problem of temporal ontology seems not of grave concern to the nefarious presentist. Although Matt Farr suggests that the question of temporal ontology in presentism can be addressed by admitting an entirely different opinion from presentism (that is, eternalism), his view makes clear the problem that beset the nefarious presentist.[11] For Farr, the discourse on temporal ontology is faced with the problem of triviality as Matt Farr notes.[12] It is quite obvious that Jonathan Tallant and David Ingram intend to avoid this problem headlong by prescribing that all presentists should be nefarious presentists.[13] Thus, the question arises on how to ground the truth of statements that exhibit spatio-temporal ontology and parts about events. For both the upstanding presentists and the nefarious presentists, the truth of presently existing events is taken as granted, given their basic claim that only present things exist. However, the crux of the matter is how to put forward the truth of past and future events in the presentists' language, such that it is backed by the truthmaker which can be referred to actual existence and at the same time can be quantified in the present. That is, grounding the truths of statements where the truthmaker is not present but can be justifiably referred to as giving support to the language of the present. Keller succinctly puts the problem thus:

---

[7] Theodore Sider, "Quantifiers and Temporal Ontology," *Mind* 115, no. 457 (2006): 79.

[8] Tallant, and Ingram, "Nefarious Presentism," 356-357.

[9] Ibid., 356.

[10] Ibid.

[11] Matt Farr, "Temporal Ontology in Perspective," accessed May 11, 2017, http://www.mattfarr.co.uk/files/theme/top.pdf.

[12] Ibid.

[13] Tallant, and Ingram, "Nefarious Presentism," 356-357.

> The problem is how there can be past- and future tensed truths though the past and future do not exist [...] the presentist can avoid the problem only if she endorses some very controversial metaphysical views. While the truthmaking problem (as I will call it) does not stand as a refutation of presentism, it does show that presentism comes at a price. Whether or not presentism is plausible depends upon whether or not that price is worth paying.[14]

We can say that the price to be paid that Keller alludes to is what Tallant and Ingram point at, which is, the jettisoning of ontological commitment by the presentist. This view is mostly a non-commitment to the temporal ontology of events that are not presently existing but are rendered in the presentists' language.

> The range of responses offered by presentists to the truthmaker objection suggests another possible response. Some presentists look to deploy "in virtue of language," without making any ontological commitments. They deny the premise (All truths require truthmakers i.e. "ontological grounds"). They tell us that there were dinosaurs is true because there were dinosaurs, where that claim is not one that commits us to the existence of anything in the past or present. More, they say that there were dinosaurs is true in virtue of there having been dinosaurs. Following a position I've argued with David Ingram call this a "nefarious" response to the truthmaker objection. If that move can be made in response to the truthmaker objection (a significant "if").[15]

The present tense quantification of a non-present event where the truthmaker is absent summarily presupposes the truth of the proposition about the event. Also, it justifies the temporal existence of the event. It does seem, however, that the position Tallant and Ingram take, gives a specific focus on the suggestion by Keller on the problem of ontological commitment as associated with truthmaking. For Keller,

> The truthmaking problem does not refute presentism, but it does leave the presentist with the twin burdens of choosing an

[14] Keller, 85.

[15] Jonathan Tallant, "Presentism, Persistence and Trans-temporal Dependence," *Philosophical Studies* 175 (2018): 2214.

> account of what underlies past and future-tensed truths and of showing that it is worth making the unattractive commitments that such an account will inevitably involve.[16]

Regarding the above statement, nefarious presentists will tend towards providing an account of truthmakers for past and future events as an unattractive commitment, as events as such are avoidable. As a philosophical system, truthmaking entails existence. The truth of a proposition entails that there must be a state of affairs that obtains or occurs as stated in the proposition. This means that what is true depends upon what exists.[17] The possibility of having truthmakers that will justify the truth of past and future events for the nefarious presentist will be a futile effort. For instance, the truth of the proposition "Julius Caesar was a Roman Emperor" will require that there is an actual period when it was described as "Julius Caesar is the Roman Emperor." However, that spatio-temporal ontology is no more captured in a state of affairs that presently exists, hence the nefarious presentists' claim that there need not be an analysis of the past event in such a manner that will require analysis by truthmaking appears convincing. This shield from analysis, that nefarious presentists allow, seems to solve the problem of plural tensed statement and quantifiers. Properties of things do exist as they are being instantiated in the state of affairs that are experienced simply. It is the attempt to find an interconnectedness among various state of affairs and properties identified that make analysis challenging.

However, these properties are somewhat undeniable aspects of reality. Heil describes properties of things as "features of the world that make a difference in how objects behave or would behave."[18] In other words, the properties that are instantiated have a way of establishing the existence of an event in spatio-temporality. These properties, according to Heil, are such as science "teases out." They are what "figures in the laws of nature, and laws govern the behaviour of objects." As much as he notes that defining the properties of things as they are instantiated is problematic, it is evident that he regards properties as dispositional powers in their possessors.[19] These properties could be referred to as causal or relational powers in the objects expressed in their interaction with the environment. Hence, for presentism, the properties of object exhibited are taken to be the truthmakers for the state of affairs in spatio-temporal ontology. There is a circular problem in this

---

[16] Tallant, and Ingram, "Nefarious Presentism," 361.

[17] Ibid., 361.

[18] John Heil, "Properties and Powers," *Oxford Studies in Metaphysics*, vol. 1, 223-254, ed. Dean W. Zimmerman (Oxford: Oxford University Press, 2004), 222.

[19] Ibid., 225.

view: if time is conceived "as a series of instants that are made out of events tied together by a relation of simultaneity, a different one for all instants,"[20] it implies that the properties of objects or events, as the case may be, are assumed in the spatio-temporal ontology where the truthmakers themselves have their source. By alluding to these temporal properties as an indicator of truthmakers, the ontology of non-present things is rendered as present truths. Still, it is in providing such a justification for an ontological commitment that the nefarious presentist turns away. Nefarious presentists argue that temporal properties do not suffice to provide truthmakers for the truth of non-present events.

According to Tallant and Ingram, the instantiation of temporal distributional properties does not admit of change as an existential fact.[21] By this, they imply that things in the past can be other than believed to be, or that which obtains. This argument is itself problematic considering that the events shift into the temporal part regarded as past are fixed. So also, when a temporal distributional property is instantiated, it is fixed and immutable.[22] This means that for an event that exists presently, its temporal distributional property – as at that point is fixed and immutable, its spatio-temporal ontology, cannot be revised. When a thing or an event instantiates a property at time T1, that instantiation itself is a truthmaker, because it is established as a fact that is actual, the possibility of a change of that instantiation as at T1 is next to zero. Whatever change will occur about what is instantiated, will be fixed in another temporal location. This includes all properties involved, either they are causal or relational.

Temporal distributional properties (of being past, being present or being future) are introduced to address this challenge. For Ross Cameron, these TDPs are capable of change such that the instantiations in the instants of time do not in any way affect the grounding of truth they provide for entities.[23] For instance, the property of age is correlative to the instantiation of TDPs, as of being-a-boy-then-being-a-man-then-being-an-old-man. The claim evident in Cameron's argument is that movement (change) of along the line of TDPs, is also exhibited in the movement (change) in the property of age. Tallant and Ingram analyze this argument as a solution to the

[20] Jiri Benovsky, "Relational and Substantival Ontologies, and the Nature and the Role of Primitives in Ontological Theories," *Erkenntnis: An International Journal of Analytic Philosophy* 73, no. 1 (2010): 107.

[21] Tallant, and Ingram, "Nefarious Presentism," 358.

[22] Ibid., 359-360.

[23] Ross Cameron, "Truthmaking for Presentists," in *Oxford Studies in Metaphysics*, vol. 6, 55-100, eds. Karen Bennet, and Dean W. Zimmerman (Oxford: Oxford University Press, 2011), 58.

truthmaker problem, which will defeat the non-commitment of nefarious presentists to ontology:

> To see how TDPs solve the truthmaker problem, consider the following (vastly simplified) example. Suppose that Ross is now a man and was a boy. "Ross was a boy" is true, but what makes it true? In Cameron's metaphysic, Ross instantiates the TDP, being-a-boy-then-being-a-man-then-being-an-old-man, and the age property. Notice that between them, the TDP and age property fix not only how Ross is now, but also how Ross was. Namely: a boy. Hence, the union of the TDP and the age property serves to make true "Ross was a boy," and so we have a solution to the truthmaker problem.[24]

However, the idea of change, which was expressed by Tallant and Ingram is overly extravagant in addition to being vastly simplified, as suggested above. This is because it can be argued that the entity "man" has properties that may not be captured by the TDP and age property; properties that may be non-spatial and non-temporal that may be of essential instantiation to the beingness of man such as rationality, moods etc. The being-a-boy-then-being-a-man-then-being-an-old-man TDP expresses a change in time and distributional property of the entity "man." This, however, does not exclude "man" from the category of entities in which it is examined through the analysis of spatio-temporal ontology and truthmaking. Non-human objects, activities and events may well exhibit TDPs, that sometimes may be confined to the scope of space.

For instance, a cup placed on a table will maintain the same spatio-temporal status except there is a change of events in its space-time boundary, such as being moved by another entity. In such case statement of fact about the cup changes. This is unlike the being-a-boy-being-a-man-being-an-old-man idea of the TDP. But for Tallant and Ingram to posit the change in the property of entities or entities themselves is to fault the need of truthmakers for a statement about the present,[25] since presently existing entities will have to remain so and as such. TDPs are of no consequence to the truthmaker problem in presentism, especially upstanding presentism. This will imply that presently existing entities are fixed in their spatio-temporal ontology, and this leads to the problem of triviality again. The statement "Phillip is the King of France now" with a state of affairs that is a truthmaker in the present, will remain fixed in the truth of the spatio-temporal ontology captured.

---

[24] Tallant, and Ingram, "Nefarious Presentism," 359.

[25] Ibid.

The concept of hyper-change and hyper-time as put forward by Ross Cameron does little to address the challenge posed against presentism as regards the idea of change. The question of time and change as being jostled by the presentists is not straightened yet, and then the concept of hyper-change and hyper-time which is pulled into the debate, only makes it more problematic.

As simple as it may sound, Tallant and Ingram conflate the idea of change and temporal location in their argument. A fixed event in the past is so, not because of the truthmaker involved, but because of the temporal location of its occurrence. The statement, "Phillip is the King of France now" will be banally false when rendered in the present where the truthmaker is fixed in the past. Furthermore, "Phillip was the King of France" will be problematic statement to analyze in reference to its spatio-temporal ontology because the truthmaker of these statements is not presently existing. Nefarious presentists avoid this problem due to the rejection of the distribution of properties over space and time. All the nefarious presentists will claim that "Phillip was the King of France" is fixed in the past as true, regarding the state of affairs that obtains for the statement to be to true. There is no strict necessitation between a proposition and the truthmaker in the present, when the temporal location of events is distributed beyond the scope of occurrence.[26]

Tallant and Ingram are of the view that presentists with a commitment to spatio-temporal ontology use necessitation loosely. For them,

> It's not merely the case that Caesar's existence necessitates the future truth of "Caesar existed." There is more to it than that. After all, the existence of any contingent existent (e.g. Queen Elizabeth's left eyebrow, Pokemon, batteries, etc.) necessitates the future truth of "2+2=4." But the relationship between Queen Elizabeth's eyebrow and the truth of "2+2=4" is quite different from the way in which the existence of Caesar fixes that it will be the case Caesar existed (and "Caesar existed" will be true). Mere necessitation seems quite the wrong way to describe the nature of this fixing.[27]

The nature of fixing here cannot be one of mere necessitation but of strict necessitation. This allows the claim by nefarious presentists that truths regarding the past do not require an analysis of the present. However, strict necessitation, as they are true cannot apply because logical truths have no temporal boundaries. The truth of the arithmetic statement 2+2=4 is

[26] Ibid., 361.

[27] Ibid.

not bounded in a spatio-temporal ontology; its truth is fixed, while the contingent existent of Queen Elizabeth's eyebrow is not fixed across time. We can argue that the relationship between the two objects will be of implication, not necessitation. That is, affirming that the truthmaker of a contingent existent is fixed, allows us to infer the necessity of logical truth in the future. However, the logical truth itself is not restricted by the scope of time.

This implies that the contingency of non-present events should allow for strict necessity; this is because the nature of fixing for the nefarious presentist is such that logical truths exhibit. Therefore, there is an interchange of the idea of necessitation and implication by the nefarious presentist.[28] The upstanding presentist may argue that implication suffices to explain how the truth of non-present events is established by the fixed properties they instantiate. But this will not do well to address the idea of "fixing as necessitation" by the nefarious presentist. If presentists generally conceive of necessitation ab initio as strict as in logical necessity, the claim of upstanding presentists fails, since the truthmakers that will admit of the truth of a proposition about the present are of strict contingency. This is because if the upstanding presentist is to go the way of primitivism where tensed truths are unanalyzed,[29] they might as well allow that mere necessity will suffice to provide an indication of truthmakers of the past statements, where truth about the past fixes or implies a truth about the present.

It, therefore, seems upstanding presentism does so much as to accommodate all experience in existence as regards temporal ontology and temporal location. But the attempt to give an inclusive existent to the spatio-temporal category of the past about truthmakers drives presentism to the fundamental question; that is how can the propositions of the non-present be regarded as true given by the truthmakers for such events which are also non-present? The result is that the upstanding presentism is inevitable as long as their position is to retain any attraction. The denial of ontological commitment in nefarious presentism implies that propositions about the present can be made in terms that will not involve triviality or banality. But will this denial by nefarious presentists cover tensed propositions about the present as well as tenseless propositions? An assumption that comes with nefarious presentism is that all talk about the present is of a tenseless kind. In other words, the un-analyzability of statements without a commitment to ontology avoids the problems of triviality, but this is a direct plunge into the principles of primitivism. The truth of propositions about the present

[28] Ibid.

[29] Ibid., 356.

need not be analyzed with the availability of truthmakers. If nefarious presentists hold this claim, then they also face the challenges that plague primitivism.

III. The Un-analyzable Present, Thisness and Triviality

Nefarious presentism as an adequate theory of temporal existence should be able to account for all the use of tenses in language. In other words, both tensed and tenseless quantification should be well explicated with the criteria that the nefarious presentists set out into their proposal. It should also be able to cover any range of quantification, either singular or plural. In combating the challenge of the truthmaker objection against presentism, nefarious presentists seem to have found a way to deny ontological commitment. A statement such as "There is a king named Charles" is understood in the tenseless sense may not be much of a problem to the nefarious presentist, even so, it faces the triviality problem. But by the extent of language use, will the appeal to un-analyzable truths explain tensed quantifiers, as in the instance of "There have been two kings named Charles?"

The principles that Tallant and Ingram endorse is that nefarious presentism should hold its consistent with primitivism.[30] Primitivism holds that the truth about the events of the past, which are within the present in their scope of utterance, are simply un-analyzable. The move to analyze such statement will only be futile, since truthmakers for such entities are not located within the temporal scope of utterance. David Lewis explains two variants of primitivism as follows:

> Version I. To be a quantifier is to function semantically like a quantifier. There must be a domain of entities, there must be a way for members of that domain to satisfy predicate phrases, and a quantifier phrase indicates whether some, or all, or none, or two, or infinitely many, or several, or [...] things in the indicated domain satisfy a predicate. Since there are no domains of past or future things, "tensed quantifiers" are not really quantifiers.
> Version II. To be a quantifier is to function inferentially like a quantifier. Tensed quantifiers are indeed quantifiers, because they obey (appropriately tensed forms of) the usual rules of quantificational logic. However, the usual semantic story about domains and satisfaction does not apply to them.[31]

[30] Ibid., 361.

[31] David Lewis, "Tensed Quantifiers," in *Oxford Studies in Metaphysics*, vol. 1, 3-14, ed. Dean W. Zimmerman (Oxford: Oxford University Press, 2004), 12.

From Version I, tenseless statements are explicated semantically without the restriction of quantifiers. In other words, the quantification of spatio-temporal entities is the mere use of words. In Version II, quantifiers are inferential. They are intended to appear in the explication of spatio-temporal entities. Although Lewis argues that the distinction between these two versions is superficial, I think it is a genuine one. If nefarious presentists are to pick up the primitivists' ambition of un-analyzability, especially the Version I., then it is evident in the move by Tallant and Ingram in laying out two principles of primitivism:

> (1) Truths about the past are expressed using primitive (and analyzable) tense operator.
> (2) The primitive (and unanalyzable) tense operators do not pick out some distinctive ontological category or aspect of reality.[32]

These two principles align with the view of primitivism as explicated by Lewis,[33] since they do not deny that tense operators indicate that there is a scope of time that needs explication in the presentists language even if the truthmaker is not present. However, as he notes, primitive tensed quantification requires a temporal analysis which the nefarious presentist will not be inclined to carry out, since this will run him into identifying temporal ontology of the past.

Thus, the nefarious presentist is left to give a presentist translation of non-present truths with recourse to triviality (in this case where the truthmaker is presently existing). Ordinarily, the problem of triviality, as presented against the presentists, can be addressed with truthmakers for presently existing entities. For instance, one can argue that when the presentist says, "only present things exist now," "only present things exist presently" and so on, it is implied that there is a truthmaker for the propositions of the present reality to be true. However, it takes a different turn with non-present reality. The language of the nefarious presentist allows the translation of non-present things in its non-present tense. For instance, "it was the case that p is true because it was the case that p." The tense operator in this proposition "was" is presented bluntly as primitive and un-analyzed. This is a case of triviality in an obtrusive sense.

The past tensed quantifier needs analysis if it is to be rendered in a presentists' language and avoid triviality that will arise from the camp of

[32] Tallant, and Ingram, "Nefarious Presentism," 361.

[33] Lewis, 12.

the nefarious presentist.[34] The nefarious presentist disregards this need for analysis. Howsoever, it is being determined, a translation of a proposition such as "There have been a king named Charles," in a restricted domain of quantification which is the present, will have to be analyzed to arrive at being explained in the unrestricted domain of quantifications in the past or future. Consider the following supposed translations of the presentists following Lewis' explication:

> "There has been a king named John."
> (a) It has been that (There is a king named John).
> (b) There has been a king named John because there was a king named John.
> (c) There has been a king named John because there has been a king named John.

From all of these statements above, it is evident that there is a level of analysis. The statements does not require being analytic to establish the necessitation of a truthmaker that has the possibility of change, as Tallant and Ingram argue. However, proceeding to give a worthwhile presentist translation, nefarious or not, requires that the past tense quantification be somewhat explicated.

In (a) above, the span operator "it has been that," is used to explicate the temporal ontology of the original statement. In contrast, the statement is substantially given by the presentist's restricted quantification. The translation (b) is somewhat primitive but not un-analyzable, and this is due to the tense operator "there has been" is summarily implied in "there was." This translation does not necessarily include a truthmaker and it does have an indication of being trivial. I am of the view that the tense operator "there was" can be further analyzed in a way that the tense operator captures the temporal location of the event.

Tallant and Ingram's view will (for consistency) align with (c), which is presented in a primitive un-analyzable way and which presents not only an unambitious move for the nefarious presentists. but also one belaboured with ambiguity and triviality. The span operator denotes a time interval that most certainly includes the unrestricted domain of quantification. To admit this for the nefarious presentist will be to deny also the restricted domain of quantification. It can be rightly deduced that nefarious presentism will hold on to an analysis of substantival ontology in its proposal to deal with quantification.

---

[34] Ibid.

> According to the substantivalist theory of time, events and things "are located at" or "occur at" different instants of time and these instants are seen as being independent of the events or things they "contain." Time is then a substance composed of such instants. It is thus easy for this view to account for periods of time without change, that is, for periods where all change stops and the universe is "frozen" during a certain interval of time since instants are not dependent on the changes that occur at them. Similarly, the view is also typically taken to be able to accommodate periods of "empty time," that is, periods where no events occur at all. Metaphorically, time is here conceived of as a container that can contain events and things, but that is capable of not containing anything.[35]

The above conception of time is consistent with the nefarious presentists' view of spatio-temporal ontology as not admitting of distributional properties. In other words, events are fixed in the spatio-temporal location in which they occur, and this suffices to explain the truth of past events at the time of utterance, that is, in the present. Furthermore, Jiri Benovsky claims that,

> [...] instants are primitively numerically distinct entities that do not require to be distinguished qualitatively (since, in the first place, they do not have a qualitative nature such that they could be distinguished in a qualitative way). Metaphorically speaking, instants are containers that are in themselves qualitatively indistinguishable and that, during a global freeze, contain qualitatively indistinguishable stuff, but that are primitively numerically distinct. Instants conceived of in such a way are what I will call ''problem-solvers.'' In short, a problem-solver is a primitive of a theory that allows us to solve a problem. In general, it is probably the case that all primitives are, at least to some extent, problem-solvers, for primitives, are typically introduced in any theory to do an explanatory job that they manage to do by having the primitive capacity to do so. In the case of substantivalism, how can the theory account for the possibility of periods of time without change? By using its primitive notion of instants that are qualitatively indistinguishable but that are primitively numerically distinct. This premise is thus a ''problem-

[35] Benovsky, 104.

> solver'' since without appealing to it the theory would not be able to accommodate the possibility of global freezes and since it succeeds to achieve its end only in virtue of the postulation that it can do so. It may seem from what I just said that I take substantivalism to be an unappealing view that does not know better than primitively postulating solutions to its theoretical challenges. But this is not so: every theory has its primitives [...] and every theory has the right to do so since without the use of such problem solvers we would not be able to get very far in the construction of metaphysical theories [...].[36]

This view of primitives that were expressed by Benovsky captures the focus of nefarious presentism in turning to un-analyzable truths of a past event (understood to be qualitatively indistinguishable instants of time described above). These instants of time, especially in the case of plural tensed quantification, are qualitatively indistinguishable because of the nature of fixing earlier discussed. For a proposition such as, "There have been five kings named Charles," as at the time of utterance, the instants of time for the temporal distribution is qualitatively absent or indistinguishable since no truthmaker(s) will ever be alluded to for the truth of the statement. This need not be of any ontological commitment to the truth of the statement, since the quantification is understood as indistinguishable in its temporal location This view tends to be consistent with what Ingram calls a "thisness ontology."[37] Ingram defines the notion of thisness as follows:

> it is a particular, primitive, purely non-qualitative property of an object; the property of being a certain object. For a given x, x's thisness is the property being-x or being-identical-with-x. For example, Barack Obama's "thisness" is the property being-identical-with-Obama. This basic idea can be sketched in various ways, but my concern is to develop a notion of "thisness" that will vindicate presentism.[38]

Making an assumption of the property of an object for Ingram is somewhat dependent on the spatio-temporality of an object. But the notion of thisness goes beyond the truthmaker's condition to establishing the state of affairs. This notion of Ingram's in order to vindicate presentism (nefarious)

---

[36] Ibid.

[37] David Ingram, "Thisnesses, Propositions and Truth," *Pacific Philosophical Quarterly* 99, no. 3 (2016): 2.

[38] Ibid., 3-4.

is clearly shifting ground from the non-commitment to ontology of nefarious presentism. The property being-x or being-identical-with-x, can be distinguished as an ontological category. The thisness of an object being described as primitive and purely non-qualitative can be regarded as implying that the property of the object is fixed. As such, the nature of fixing that Tallant and Ingram argue for plays out in the thisness of the present.

> (a) The Nature of a Thisness – every object instantiates a thisness; a thisness is a particular, primitive, purely non-qualitative property of an object.
> (b) The Life of a Thisness – a thisness T comes into being with an object x, T is uniquely instantiated by x throughout x's existence, and T continues to exist uninstantiated when x has ceased to exist.
> (c) The Character of a Thisness – a thisness T instantiates higher-order properties, which characterizes the object that instantiates T, x; the relevant properties of T initially correspond to the lower-order properties of x.[39]

From the above, the thisness of an object continues to exist even when the truthmaker of the object is no longer present. This is a quite ambivalent position to hold for a nefarious presentist, and as such the coherence of the principles (un-analysability, non-qualitativeness) of the notion of primitivism is questionable as the nature, life and character of the thisness as a property of an object do not solve the truthmaker problem pointedly. How is thisness so different from the temporal distributional properties that Cameron suggests? Phil Corkum notes that Tallant and Ingram are caught into the web of temporal extension beyond the present and the legitimacy of the acceptance or denial of this extension is yet to be sufficiently answered by these scholars.

> Tallant and Ingram thus take the distributionalist to be caught in an inconsistency: temporal distributions require extensions of time, but presentists must deny that there are such extensions.[40]

Contrariwise, the idea of thisness as a property of objects in spatio-temporal ontology as Ingram put forward seems to reintroduce the idea of distributional property into presentism. How can we possibly describe properties (as a

[39] Ibid., 5.

[40] Phil Corkum, "Presentism, Truthmakers and Distributional Properties," *Synthese* 191 (2014): 3433.

distribution or thisness)? Corkum discusses the legitimacy of reference to temporal distributional properties a truthmakers:

> Are temporal distributional properties legitimate truthmakers? Under what conditions do such properties conform to the prohibition against cheating? [...] temporal distributional properties meet this requirement just in case they are not reducible to non-distributional or uniformly distributional properties. For readability, let me say that temporal distributional properties are reducible or irreducible, and leave the reference to that to which they are or are not reducible, non-distributional or uniformly distributional properties, tacit.[41]

Corkum's argument above also suggests the un-analyzability of the present, even with the distribution of properties as the attempt by Ingram to use thisness does little to address the analysis of the present without the distribution of properties of the object beyond the present. Hence, the presentist claims about the spatio-temporal location of an object when the truthmaker is no longer in the present can well be acknowledged. This move towards primitivism by nefarious presentists also faces the problems of semantics and inference, as indicated by Lewis. However, with the notion of thisness and distribution of property, once an object instantiates the property in existence, it addresses the tensed quantification in the domain of entities under semantics and also the quantifier phrases in inference. Even so, it may well attempt to reduce the demand of analyzing span operators, which, as Lewis argues, cannot be properly understood when they are un-analyzed.

If nefarious presentism admits that objects which are into the spatio-temporal location have properties and that there is a nature of fixing on those properties which are un-analyzable without losing their spatio-temporal location, then the burden of triviality is also on the nefarious presentist, if an attempt of explication is made. The statement "There was a king named Charles in Britain" will be accepted by nefarious presentists not due to its ontological commitment to a truthmaker that supervenes on the beingness of a state of affairs, but due to the spatio-temporal location which is a present T1, when it did exist that "There is a king named Charles." The properties of the entity are regarded as distributed over time, as the present is no longer existing in its spatio-temporal occurrence, and by implication for the nefarious presentist it is un-analyzable.

[41] Corkum, 3434.

IV. Conclusion

The nefarious presentist successfully avoids to analyze these statements about the present. The primitive, in this case, will be either tensed or tenseless quantifiers, span operators, in singular or plural quantification. Tallant also notes that the nefarious presentists' position allows them to be described as "cheats" without an ontological commitment,[42] and they claim to provide a sufficient presentist view on the un-analyzability of the present. Benovsky also makes the claim that primitives in a way cut off the discourse on a topic too soon for a conclusion.[43] If an object or property or word can be considered as un-analyzable, the need to provide further explication on the proposer is somewhat reduced. This is expressly shown in the following example:

> nefarious cheats (like Tallant) will say "it was the case that p" is true because it was the case that p'. Here, the "was" is primitive and unanalyzable. Notice, also, that Tallant will agree that this "was" does not help us to pick out any ontology; it does not help us to speak of any distinctive metaphysical category or kind.[44]

The un-analyzability of the tense operator 'was' will inevitably lead to the problem of triviality, if the distribution of properties over spatio-temporal locations is admitted by the nefarious presentist, even in the mildest sense. as it is evident in the notion of thisness, that Ingram suggests, or in virtue of a property having existed, as Tallant suggests.[45] It is quite evident that although nefarious presentists claim that their view does not pick out any ontology, the insistence on this claim will end up being a self-defeating one. A tense operator applies to a thing, either concept or object, and the existence of it, even of the Menoingian kind can be inferred. Hence the nefarious presentists will face the challenge of triviality of truth in their claim of un-analyzability.

References

Benovsky, Jiri. "Relational and Substantival Ontologies, and the Nature and the Role of Primitives in Ontological Theories." *Erkenntnis: An International Journal of Analytic Philosophy* 73, no. 1 (2010): 101-121.

---

[42] Tallant, and Ingram, "Nefarious Presentism," 361.

[43] Benovsky, 106.

[44] Tallant, and Ingram, "Nefarious Presentism," 363.

[45] Tallant, "Presentism, Persistence and Trans-temporal Dependence," 2214-2216.

Cameron, Ross. "Truthmaking for Presentists." In *Oxford Studies in Metaphysics*, vol. 6, 55-100. Edited by Karen Bennet, and Dean W. Zimmerman. Oxford: Oxford University Press, 2011.

Corkum, Phil. "Presentism, Truthmakers and Distributional Properties." *Synthese* 191 (2014): 3427-3446.

Crisp, Thomas M. "On Presentism and Triviality." In *Oxford Studies in Metaphysics*, vol. 1, 15-20. Edited by Dean W. Zimmerman. Oxford: Oxford University Press, 2004.

Farr, Matt. "Temporal Ontology in Perspective." Accessed May 11, 2017. http://www.mattfarr.co.uk/files/theme/top.pdf.

Heil, John. "Properties and Powers." In *Oxford Studies in Metaphysics*, vol. 1, 223-254. Edited by Dean W. Zimmerman. Oxford: Oxford University Press, 2004.

Ingram, David. "Thisnesses, Propositions and Truth." *Pacific Philosophical Quarterly* 3 (2016): 442-463.

Keller, Simon. "Presentism and Truthmaking." In *Oxford Studies in Metaphysics*, vol. 1, 83-104. Edited by Dean W. Zimmerman. Oxford: Oxford University Press, 2004.

McKinnon, Neil. "Characterising Presentism." In *New Papers on the Present: Focus on Presentism*, 13-30. Edited by Roberto Ciuni, Kristie Miller, and Giuliano Torrengo. München: Philosphia, 2013.

Lewis, David. "Tensed Quantifiers." In *Oxford Studies in Metaphysics*, vol. 1, 3-14. Edited by Dean W. Zimmerman. Oxford: Oxford University Press, 2004.

Sider, Theodore. "Quantifiers and Temporal Ontology." *Mind* 115, no. 497 (2006): 75-97.

Tallant, Jonathan, and David Ingram. "Nefarious Presentism." *The Philosophical Quarterly* 65, no. 260 (2015): 355-371.

Tallant, Jonathan. "Presentism, Persistence and Trans-temporal Dependence." *Philosophical Studies* 175 (2018): 2214-2216.

G. Elfstrom · *Conatus* 5, no. 1 (2020): 27-53
DOI: http://dx.doi.org/10.12681/cjp.21995

The Theft: An Analysis of Moral Agency

Gerard Elfstrom
Auburn University, USA
E-mail address: elfstga@auburn.edu
ORCID ID: https://orcid.org/0000-0002-5604-6556

Abstract

Adam and Eve's theft marks the beginning of the human career as moral agents. This article will examine the assumptions underlying the notion of moral agency from the perspective of three unremarkable human beings who found themselves in situations of moral difficulty. The article will conclude that these three people could not have acted differently than they did. It will conclude that it is unreasonable to assume that ordinary human beings will inevitably possess the resources to address difficult moral decisions.

Key-words: *moral agency; moral responsibility; moral luck; free will; Holocaust; Ford Pinto case*

We began badly. Adam and Eve were commanded to avoid eating fruit from the tree of knowledge of good and evil. Like wayward children, they were nonetheless tempted and succumbed to temptation. And, like wayward children from their day to ours, they hid in shame when their transgression was discovered, then tried to shuffle off blame. Their childish ruses failed to save them from God's wrath, and they were banished from the Garden of Eden.

Genesis intrigues and perplexes. It intrigues because its main characters are patently human. This ancient narrative describes behavior that remains familiar to every household. It perplexes because fruit theft is hardly a major crime. Adam and Eve's fatal error was disobedience to divine command, but why did the deity guard that particular fruit so vehemently? Children routinely disobey their parents and are routinely punished for doing so, but a sentence of hard labor, pain, and exile from paradise seems excessive even by

Old Testament standards. An aside has a clue to the deity's ire. He laments, "Behold, the man has become like one of us, knowing good and evil."[1] Moral understanding made us godlike. Apparently the deity had an additional fear, "And now, lest he put forth his hand and take also of the tree of life, and eat, and live for ever [...]"[2] Moral understanding is one step removed from eternal life and thus full divinity. Only timely banishment from the Garden of Eden kept us from immortality and membership in the company of the gods. So, it seems that Adam and Eve's theft aroused the deity's fury because it wished to bar them from knowledge of good and evil. That knowledge is both the curse and the glory of humanity; its possession drove us from paradise, but its possession makes us godlike. As Prometheus' theft of fire marked the beginning of human technological prowess, the theft of moral understanding marked the beginning of our fitful career as moral agents, that is, as individuals capable of being morally guilty or morally praiseworthy.

Hasty intervention denied us eternal life, but we retain moral understanding. So far as we can discern, no other beings possess moral understanding, and therefore no other creatures are capable of being morally virtuous or vicious. Though necessary for moral agency, moral understanding by itself does not suffice to make us morally accountable. Adam and Eve became sinners because they were capable of acting in accord with the deity's command but did not. Before they gained moral understanding, they were able to perform the acts they desired and gain the results they sought. They could act as they pleased, and, in particular, they could be obedient or disobedient, an ability shared with dogs, cats, and human infants. Their control over their actions allowed them to disobey the deity's command and become sinners. But, this particular sin, the theft of moral understanding, transformed Adam and Eve; they were now moral agents and able to add moral guilt to their guilt of disobedience. Just as individual infants become morally accountable only when they add moral understanding to their ability to obey or disobey, the human species became morally accountable when it added moral understanding to its ability to act as it wished. These three qualities, moral understanding, the ability to choose the principles that guide our actions, the ability to choose our actions, and the ability to gain the ends we seek comprise the essential features of moral agency.

Though Adam and Eve's humanity remains familiar, the conditions that shaped their lives are long gone. Their post-Eden world of small flocks, compact fields, and small villages has been replaced by another that is vastly different and offers starkly different moral challenges. History has worked

[1] *Genesis*, 3:22.

[2] Ibid.

unsettling alchemy on our loot. Adam and Eve became guilty because they stole moral understanding. We moderns are more likely to falter through lapses in moral understanding, constraints on our actions, confusion over when we are obligated to exercise moral agency or how we should do so. The century just past offers striking examples of individuals whose lives disturb precisely because the conditions of their moral agency were imperfect. Our ancient theft made us moral agents, but we cannot evade the accountability moral understanding brings, not even when the elements of our moral agency have been compromised. Loosing hold of our booty will not return us to the Garden of Eden or to innocence. Human nature has not changed. Rather, the conditions of human life have changed. Consider Wilhelm Trapp.

I. The case of Wilhelm Trapp

Wilhelm Trapp was a decent man, a respected and honest member of his community. He fought honorably in the World War I and was decorated for his efforts. He was concerned for the well-being of those around him, was sensitive to their feelings, and routinely made small efforts to assist them. He was known as "Papa Trapp."[3] Though he had conducted himself honorably as a soldier, he hated to see people suffer and was viscerally repelled by the thought of killing human beings.[4]

Trapp was also a war criminal, guilty of the gravest crimes against humanity. He was commander of Reserve Police Battalion 101, a unit of the Nazi German reserve police forces charged with exterminating Jewish communities in Poland during the World War II. Major Trapp commanded the slaughter of thousands of innocent people, including small children and infirm elderly. Repeatedly he ordered his men to pull innocent people from their homes and kill them *en masse* at point blank range. By all accounts he hated giving this command, and he avoided being present at the scenes of slaughter when he could.[5] He also allowed the men under his command to evade the murder if they wished.[6] He never threatened to punish those who sought to avoid killing, and he never cajoled any of his men to commit murder. Several members of his brigade thought him weak and cowardly because he was repelled by the massacres he ordered them to carry out.[7]

---

[3] Christopher R. Browning, *Ordinary Men: Reserve Police Battalion 101 and the Final Solution in Poland* (New York: Harper Collins, 1992), 2, and 44-45.

[4] Ibid., 2, and 57-58; Daniel Goldhagen, *Hitler's Willing Executioners* (New York: Knopf, 1996), 212-213, 240, and 537, n 22.

[5] Browning, 57-58.

[6] Browning, 2, 57, 86-87, 102-103, 130, and 171; Goldhagen, 213-215.

[7] Browning, 45-46, 56, and 57-58.

Major Trapp was not an evil man, at least not before he entered Poland and maybe never. Had the war not intruded, he would doubtless have spent his life a conscientious and respected citizen of Battalion 101's home in Hamburg, Germany. In death, he would likely have been fondly remembered by the few who were aware of his existence. He would have ranked with the obscure legions of the world's simple, ordinary, upright people, the salt of the earth whom anyone would be pleased to have as a neighbor. There is no evidence that his character changed in any marked way during the war. He remained the sober, conscientious official he had been all his life, yet he also became a war criminal and repeatedly ordered his men to commit actions that sickened him.[8] Though his character was not evil, the acts he ordered certainly were, and the vast killing machine that swept him along certainly was. Trapp may well have insisted that the commands he gave, though he gave them repeatedly and ensured they were carried out, were not his. In two senses, they clearly were not. Apparently, he had no desire to give them, and, left to his own devices, he would never have done so. They are not the commands a man of his character desired or valued. He was bitterly repelled by them and the suffering they brought to innocent people.

Yet, though he was decent, sensitive, and honest, it apparently never occurred to Trapp that he could refuse to issue his commands. He had, after all, been ordered to give them. He seemed to be unaware that he had any other choice, and that is part of the explanation of why he would likely have claimed the commands were not his. Also, it apparently never occurred to him that the murders he ordered were morally wrong. Though he was repelled and emotionally shattered by them, he never drew the conclusion that they were morally abhorrent. It seems a simple and direct inference, but Trapp never made it.[9]

---

[8] With some consternation, Goldhagen cites one of Trapp's men. Speaking some years after the war, the Battalion 101 veteran observed, "He was what one would call a fine human being and I deem it impossible that it was he who had ordered the shooting of the hostages." Goldhagen then comments, "Trapp – who years later, despite having led his men in mass murder, is remarkably pronounced 'a fine human being'[...]" Goldhagen, 240. Goldhagen's bemusement captures the difficulty perfectly; Trapp was at once a decent human being and a mass murderer.

[9] Browning and Goldhagen, though they employ the same sources and are alert to the limitations of their sources, disagree on the degree of Trapp's anti-Semitism and the degree to which he eventually became enthusiastic about his assignment. As previous references show, both agree that he was widely regarded as a kindly person who initially was greatly troubled by his assignment and continued to have conflicting emotions about it. They also agree that he eventually lost at least some of his inhibitions and came to display some enthusiasm for his task. Browning, however, believes the anti-Semitic prefaces Trapp gave his orders to kill were unfeeling transmissions of official doctrine used to make an unsavory task easier for his men (Browning, 102 and 149). Goldhagen believes this is mistaken, that Trapp's anti-Semitism was genuine and that he took satisfaction from the killing, or, at least, that he eventually came to do so (Goldhagen, 550-551, n. 61). If Goldhagen is correct, Trapp's racism may

It would seem that Major Trapp could have avoided his torment by simply refusing to carry out the orders he received. On reflection though, it is not obvious that he was capable of this. A stronger, more self-assured individual with a different view of authority may well have refused to transmit the orders Trapp received from above. A stronger individual may have done more than refuse, may instead have actively fought the Nazi machine. But, it's not obvious that Trapp, given his character, background, and values, could have acted other than he did. Even if he had somehow found the resources to battle against the program of extermination, he would have been crushed, and his effort would have been entirely fruitless. He was an ordinary man. A saint or a hero might have risen to noble but futile self-sacrifice, but Trapp was neither. And, neither are the great majority of the rest of us.

Apparently his breeding suited him admirably for the role of a responsible citizen in a bustling German port city. Had he remained in Hamburg's tidy surroundings, he would have led the exemplary life many simple, decent people achieve. He had been born and bred to do his duty, to follow orders, and trust in the system that sustained him. For the bulk of his life, this conditioning served him well. But, nothing in his background equipped him for the circumstances that awaited him in Poland. It is difficult to say how he would have responded if he had been ordered to kill members of his own family. Would he have obeyed that order too? Would he have refused? Would he have recognized its moral repugnancy? It is hard to know. It's not obvious that he was capable of grasping the enormous moral evil of his actions in Poland. And, it's difficult to claim that he could, in any realistic construal of the term, have acted other than he did.

If Trapp was not evil, at least not initially, it is also not easy to determine exactly where the evil lies. Battalion 101's actions were evil, that's for sure, but not all who carried them out were evil. Some, like Trapp, were

have prevented him from recognizing the immorality of his acts. However, this cannot be the complete explanation, since it is possible to be a racist without endorsing the mass slaughter of innocents. Furthermore, the fact that Trapp's enthusiasm appeared only after some months of slaughter gives support to the view that he was simply ground down by his assignment. The scant and unreliable data available to us are consistent with both Browning and Goldhagen. Hence, Trapp was either a decent man who never reconciled himself to the killing he ordered, or he was a decent man who came to embrace the killing, and therefore became an evil man – but without losing his innate decency towards human beings whom he was not ordered to kill. He never became evil through and through. The view that the magnitude of his evil is sufficient to infect his entire character and transform him into an entirely evil person is plausible. However, adopting this view would not allow a moral distinction between Trapp and those who were eager to slaughter from the first or those who were callous and brutish in all aspects of their lives. Whatever slant is given to the data available, it remains true that Trapp was a man driven by circumstance to perform acts he would have rejected if left to his own devices.

apparently sickened by their assignment and had no taste for it.[10] Others, those who embraced their assignments and carried them out enthusiastically, were evil, but, without the Nazis' intrusion, most of them would never have contemplated the actions they undertook. Several of this latter group were described by their peers as young, ambitious people who were eager to get ahead, and killed energetically because they believed it would help advance their careers as party functionaries.[11] So, they were evil, but not necessarily because they reveled in slaughter for its own sake. They were evil because they were glad to engage in mass killing when it served their personal interests.

Nonetheless, the 101 Battalion's evil and non-evil men worked together to carry out the tasks assigned to them, and the chilling implication is that the evil or non-evil nature of its men made little difference to the result.[12] Had the entire battalion been comprised of people like Trapp, it would likely have carried out its assignments all the same, without enthusiasm perhaps, and maybe less efficiently, but the job would have been done. The Nazi Holocaust machine took in shallow careerists, hangers-on, the weak, the strong, and also simple, honest, decent people, like Major Trapp, then made them killers.[13] The diverse character and values of its human instruments mattered little to its operation. The machine could not have functioned without them, but it had to twist and channel their diverse natures to serve its purposes, and it succeeded.

Major Trapp, those ordinary Germans who, like him, were decent and honest, the Holocaust, and Nazi Germany pose a direct challenge to our

---

[10] Browning, 57-58, 59, 62-63, 69, 73-75, 86-87, 102-103, 113, and 168; Goldhagen, 261.

[11] Browning, 75-76.

[12] Both Browning and Goldhagen note that the men of Battalion 101 were highly unpromising candidates for the role of mass murderers; their backgrounds, age, and social standing contained nothing to support the view that they would be capable of cold-blooded murder. Browning, 164-165; Goldhagen, 206.

[13] Different authors have varying explanations of the mechanisms which drove these unremarkable people to murder. Browning believes the deciding factor was peer pressure, the desire to avoid appearing weak in the presence of peers (Browning, 184-186, and 375-416). Goldhagen believes the Nazis were able to exploit ordinary Germans' inherent anti-Semitism to transform them into killers (Goldhagen, 13-14). Zygmunt Bauman suggests that a special code of honor of the civil servant, the ability to execute superior orders as though they were one's own and to sacrifice one's own concerns and self-interest while doing so, is a key to understanding the crime. See Zygmunt Bauman, *Modernity and the Holocaust* (New York: Cornell University Press, 1989), 21-22. Bauman also cites the view of Herbert Kelman, who claims that the triple factors of authorization from on high, routinization, and dehumanization of victims together made the killing possible (Bauman, 21). Each of the authors recognizes that a variety of plausible explanations have been offered. In fact, the plethora of explanations deepens the mystery of the crime rather than dissolves it; the ease of devising plausible explanations and the difficulty of eliminating any of them, heightens the sense that the crime lays beyond explanation.

ancient legacy of human moral agency. Trapp and his compatriots participated in one of the greatest moral outrages of human history, yet our traditional conception of moral agency falters when called upon to judge men like Trapp. It has few resources to morally assess ordinary, decent human beings when vast, complex human institutions knit their actions together with the efforts of many people to produce evil or human suffering.

Adam and Eve became sinners for violating a command which they could have obeyed. Major Trapp is guilty for obeying a command which it is unlikely he could have disobeyed.[14] He is guilty nonetheless. Adam and Eve were guilty for what they did and knew to be wrong. They were guilty for doing what they could have avoided doing. Trapp is guilty for failing to know what he should have known but likely could not know, namely, that the slaughter of innocents is morally abhorrent. He is guilty for the acts he performed, even though he could not have averted their result and he lacked the strength of character to refuse to carry them out. He is guilty, even though he brought no evil to Poland, because he should have understood the evil of his actions. He should have understood even if he was incapable of doing so – because we cannot overlook evil of this magnitude. Trapp was extradited to Poland in 1947. In 1948 he was tried, then executed for his crimes.[15] Who can say he was not guilty?

And yet, because Trapp didn't understand the evil of his act and was incapable of doing so, he lacked the personal and moral resources to refuse the act, and he could not have prevented the slaughter from occurring, we must remain uncomfortable with the judgment that he is guilty. The circumstances of his conduct and the facts of his character count against guilt. He must be guilty, but he also cannot be.

Trapp never lost his conventional morality or innate decency. Apart from the killing, he remained the man he had always been. But, he didn't recognize that he could disobey commands he had been given, and he apparently did not understand that the acts he ordered were morally wrong. Perhaps he was anti-Semite, and his racism prevented him from recognizing the wrong of

[14] Though he is careful to acknowledge the pressures on Trapp and his men, the turmoil of their circumstances, and the complexity of their motives, Browning believes that each of the men in the Battalion made the choice to kill, could have made a different decision, and, thus, each is morally accountable for his actions. He bases his conclusion on evidence that at least some of the men refused to kill, others refused after awhile, and others avoided killing when they could slink away from it. Hence, because different men behaved differently, all had the choice of how to behave. Browning, 188. Unfortunately, this does not take account of differences in the nature and personal resources of individual human beings. Some are self-confident and sure of their values, others are weak and insecure, while others are devoted only to nurturing their own interests. The fact that some, such as Martin Luther King or Andrei Sakharov, are great moral heroes is not evidence that the rest of us are capable of moral heroism.

[15] Browning, 144.

his acts. Or, perhaps he was simply incapable of understanding that orders handed down from higher authorities could be morally wrong. But, whatever the explanation, it is unlikely that Trapp was capable of recognizing the evil of this conduct or that he could have refused to perform it. So, he is guilty, but his guilt is deeply troubling.

II. The case of Otto Stange

So far as we know, Otto Stange never personally harmed anyone. We have only a few small scraps and bits of information about him, like stray beams of light filtering through closed blinds. In contrast to our vivid portrait of poor, tortured Wilhelm Trapp, Otto Stange's life remains in the shadows of his unprepossessing office. Yet, his role in the Holocaust was as critically important as Trapp's. Otto Stange was a minor bureaucrat, an *Amtsrat*, in the German railway system. He was charged with devising *Sonderzüge*, special trains.[16] *Sonderzüge* were not regularly scheduled. Rather, they were individually planned and assembled to serve particular purposes. Some were contrived to transport ethnic Germans for resettlement. Others moved the mentally ill to killing centers. Yet others transported Jews out of the Third Reich, mainly to the concentration camps and death camps in Poland.[17] Requisitions for special trains to transport Jews originated in Adolph Eichmann's bureau in the Reich Security Main Office. Eichmann was in charge of its "resettlement" section. One of Eichmann's underlings, Captain Franz Novak, was liaison to the *Reichsbahn*, the German railway system. Novak would deliver requisitions for special trains to an *Amtsrat*, often Stange, who would devise a transport program for each special train, then hand it over to Novak, who carried it to Eichmann's bureau.[18]

Stange would have known that his special trains would carry Jews, since Novak's requisitions were clearly labeled as such.[19] He must have known the *Sonderzüge*'s destinations in order to draw up his transport programs. He may or may not have known of the Jews' fates once his special trains had done their job. If he did not know, it would have been a simple matter for him to ask Novak. One scrap of information reveals that, "He was very convinced of the importance of his work and his person," another that, "He was 'choleric,' sick with gallstones, and once hospitalized."[20] In his sixties during the World

[16] Raul Hilberg, "German Railroads/Jewish Souls," *Society* 14, no. 2 (1976): 68.

[17] Ibid., 527 and 536-537.

[18] Ibid., 535.

[19] Ibid., 539, and 544.

[20] Ibid., 535.

War II, Stange had worked in the *Reichsbahn* all his life. The details of his work with Novak likely differed very little from the labor he had performed for decades. Only the ultimate result differed.

Of the Jews killed in the Holocaust, approximately half were transported to their fate by train. Many were carried by the special trains Stange planned.[21] During the early portion of World War II, the *Reichsbahn* employed nearly a million and a half people, and dispatched some 20,000 trains each day. It owned 850,000 freight cars, of which perhaps 15% were empty at any one time.[22] Novak's requests for Jewish special trains had to be balanced against the demands of military transport, industrial freight, and ordinary commercial traffic. As the war progressed the *Reichsbahn* was stretched to its limit, and strained further by Allied bombing and partisan sabotage. Nevertheless, as a result of the skills and effort of men such as Novak and Stange, the Holocaust machine always found ways to assemble trains that transported Jews to their deaths.[23]

The Holocaust could not have been carried out without the operation of sprawling organizations like the *Reichsbahn* and many millions of ordinary, unremarkable, skilled, and conscientious persons. They were persons like Otto Stange, people who employed discipline and abilities developed long before the war to meet essential needs of a modern, industrialized nation.[24] The grim arithmetic of the Holocaust drives the point home. It snuffed out the lives of nearly 6 million Jews. The animal fury unleashed in the *Kristallnächte* resulted in maybe 100 deaths. Nearly 200 years of *Kristallnächte* mob violence would have been required to kill the 6 million people the Holocaust eliminated in a few years.[25] A crime of this magnitude cannot be fueled by brute, criminal rage. It requires sprawling, carefully organized, and efficient organizations, those like the *Reichsbahn*. There was little place for the psychopathic, the deranged, or the criminal in such operations. In fact, they were carefully screened out.[26] The effort required legions of ordinary, disciplined, honest, and skilled people, people of the same sort that the world's large organizations require to this day.

The scraps and bits of information we have reveal that Stange was likely a lesser human being than Wilhelm Trapp. Stange was a sickly, self-important,

---

[21] Ibid., 543.

[22] Ibid., 536.

[23] Ibid., 532.

[24] Ibid., 548-550.

[25] John P. Sabini, and Mary Silver, "Destroying the Innocent With A Clear Conscience: A Sociopsychology of the Holocaust," in *Survivors, Victims, and Perpetrators: Essays on the Nazi Holocaust*, ed. Joel Dinsdale, 329-357 (Oxfordshire: Taylor & Francis, 1980), 329-330.

[26] Bauman, 19-20.

minor bureaucrat. Trapp cared about other people, tried to help them in many small ways, and was revolted to the core of his being by the death he brought. It is likely that Stange was placidly free of any of this torment. It's difficult to say that Stange was guilty or what his crime may have been. Tens of thousands of deaths resulted from his activity, but was he guilty of them? Should he have made himself aware of the consequences of his actions and refused to carry out his usual duties? Should he have tried to combat the Holocaust? A better man would have done so, but Stange was the man he was. It is not self-evident that Stange was morally obliged to become a better person, and it is not obvious that he was able to become one, given his lack of intelligence, fortitude, and initiative.

Despite its vital importance to the Holocaust and considerable evidence that some of its workers were viscerally aware of its role in the slaughter, no employee of the *Reichsbahn* was tried for war crimes. None was so much as summoned as a witness in the Nuremberg Trials.[27] Stange carried out his ordinary duties in his ordinary way, yet they were crucial links in a vast chain of human actions that brought mass death. During the war, he may well have walked a bit straighter, lifted his head a bit higher, and spoken with a bit more authority. He likely had a sense that he played an essential role in something that was highly important – he received frequent visits from a high ranking official after all and performed valuable services for him. His self-satisfied bearing would have been justified. But, on the other hand, if, following the war, he were questioned about his role in the mass slaughter, he would have likely responded that he only carried out his small duties as best he could and that their large and remote consequences were far beyond his responsibility or understanding. He was not authorized to concern himself with them.

And, he would have been absolutely correct to say so. But, the fact remains that the Holocaust could not have occurred without him and millions more like him. Stange assisted Adolph Eichmann's "resettlement" program. Eichmann's responsibility was only to transport Jews, to "resettle" them. Like Stange, he commanded no one to die. He was courteous to his staff and those around him. Eichmann repeatedly asserted that he was, like Stange, a mere cog in the machine.[28] Eichmann professed to have kindly feelings toward Jews and to be intrigued by their culture.[29] He visited the death camps, and was sickened by what he saw.[30]

---

[27] Hilberg, 523, and 544-545.

[28] Hannah Arendt, *Eichmann in Jerusalem: A Report on the Banality of Evil*, revised edition (New York: Penguin Classics, 1977), 289.

[29] Ibid., 26, and 40-42.

[30] Ibid., 87-88.

So both Stange and Eichmann were in the business of transporting Jews. Yet, Eichmann was guilty and paid for his guilt with his life. He was guilty for viewing himself merely as a cog, an instrument, and failing to recognize that he was a human being, with the moral responsibility other human beings possess.[31] Then how can Stange be innocent? Stange too was a gear in a large and complex machine, but neither he nor Eichmann was only a cog. Their work as cogs demanded the intelligence, understanding, and self-control that only human beings possess. The two were able to play their roles in the Holocaust only because they were human and had human abilities. Because Stange was human and possessed the abilities of other humans, he was obliged to recognize that he had the moral responsibilities other human beings possess. Nonetheless, it is not evident that he was guilty for acting as he did.

To be sure, there are crucial differences between Eichmann and Stange. Eichmann's explicit responsibility was to transport Jews to the death camps. If the railroads had become unavailable, he could have sought other means. Stange's responsibility was only to devise *Sonderzüge*. If he had ceased to receive visits from Captain Novak, he would no longer have played a role in the Holocaust. Eichmann was in charge of the entire resettlement effort. Stange played a small role in that program. Eichmann visited the death camps and knew what happened there. Stange likely did not know. Further, Eichmann, like some of Trapp's men, was a shallow careerist. He was eager to advance himself, and accepted the deadly result of his efforts as the price success in his career demanded. Stange was likely too old and too settled in his ways to be interested in getting ahead. But, it's not obvious that these differences suffice to absolve Stange from guilt, and the question remains undecided.

Stange is disturbing. He disturbs because his guilt or innocence is unsettled. Adam and Eve were guilty for what they did. Stange disturbs because the activity he had performed for his entire adult life suddenly became an essential component of one of the greatest crimes of human history, and we lack a moral framework that would allow us to determine whether he was guilty or remained innocent. Our moral thinking has few resources for assessing the moral culpability of those who contribute to evil but play minor roles bringing it about and are several steps removed from its result.

But a major part of why Stange is more disturbing than Trapp is that his life and his circumstances are so much closer to ours. Trapp was pulled from the surroundings his ordinary morality was capable of addressing and was placed in circumstances it was unable to address. Stange, on the other hand, remained in his familiar office and carried out his usual business. Only the ultimate result of his labors changed. We are unlikely to be called upon to command the death of thousands of innocent people. Yet, it is quite

[31] Ibid., 289.

possible that we may one day find that the skills, habits, and discipline we have developed over a lifetime have become crucial links in a chain of human actions that result in harm to other human beings. Worse perhaps, and more likely, the rest of us sober, hard-working, skilled people may, like Stange, continue to labor diligently in vast organizations and play crucial roles in endeavors whose results we understand only dimly and whose consequences for other human beings range beyond our comprehension. And then, we may wonder about our own guilt or innocence because we have scant resources to morally assess honest and innocuous labor that somehow becomes part of a chain of activity that brings evil to human beings.

III. The case of Dennis A. Gioia

Dennis A. Gioia is a very different person than Stange or Trapp, and he lives in a starkly different world. Certainly he is no Nazi, and he is in no way associated with horrors remotely akin to those perpetrated in the Holocaust. Yet, there are several illuminating and intriguing parallels between his life and his circumstances and those of Trapp and Stange. And, like Trapp and Stange, his career puzzles and unsettles.

Gioia was an employee of the Ford Motor Company from 1972 to 1975. Fresh from college and the counter-culture movement of the era, he had routinely demonstrated against the war in Vietnam and was an energetic critic of large corporations. He joined Ford partly, he asserts, with the idea that he could help transform it from within.[32] So, he left campus with his values intact and planned to impart them to the vast organization he entered. He was pleased to find his work challenging and interesting, and he devoted himself to his assignments with enthusiasm and a strong conviction that he could make a difference. He was soon immersed in his duties and Ford's corporate culture, his activist's long hair now cropped short.[33] One result of his diligence is that in 1973 he received the assignment of Field Recall Coordinator. The position included the responsibility to monitor information about possible safety or functional defects in Ford's products and initiate requests for recall and repair or reconfiguration of vehicles that were judged unsafe or otherwise defective. The assignment was demanding and exhilarating. He took his new responsibilities very seriously, and he was, initially at least, keenly aware that the lives and well-being of other human beings depended on his diligence.[34]

---

[32] Dennis A. Gioia, "Pinto Fires and Personal Ethics: A Script Analysis of Missed Opportunities," in *The Ford Pinto Case: A Study in Applied Ethics, Business, and Technology*, eds. Douglas Birsch, and John H. Fielder, 97-117 (New York: State University of New York Press, 1994), 98.

[33] Ibid., 99.

[34] Ibid., 99.

He inherited responsibility for 100 active cases when he took his new position, approximately half of them involved safety matters, that is, possible defects that might result in death or injury.[35] Though harried, he threw himself into his assignment and sought to do his job well. He acknowledges, however, that his early vivid sense of responsibility for human life gradually gave way to emotional numbness and professional detachment from the human suffering that might result from defective products. He was trained to scan consumer complaints and accident reports for evidence of an unusually high rate of component failure and obvious patterns of causes for these failures. If he believed he had found such patterns, he had the authority to request that his department review the case and consider issuing a recall notice.[36]

During Gioia's tenure, Ford Motor Company was busily manufacturing and selling millions of Ford Pintos, small, cheap, economy cars devised to combat the inroads on its sales inflicted by small, thrifty cars from Germany and Japan. The Pinto had become notorious by the time Ford ceased production in 1980. There were reports of fires resulting from rear-end crashes at low and moderate speeds. One victim of such a crash sued Ford and was awarded millions of dollars in damages.[37] The balance of public opinion was likely tripped against the Pinto when a counterculture magazine, *Mother Jones*, published a vigorous and lurid expose of the Pinto in 1977.[38] The report and the controversy it ignited alerted the public to the possibility of a deadly defect in the Pinto. Shortly after the *Mother Jones* piece appeared, the Federal National Highway Traffic Safety Association began an investigation. It completed its study in 1978 and concluded the Pinto was defective. In consequence, it recommended that Ford initiate a recall of 1971 through 1976 Pintos, and Ford promptly complied.[39] The problem is that its gas tank was located between the rear axle and the rear bumper, an unprotected location. In the event of an impact from behind, the tank could be rammed into the rear axle and differential which was located in the middle of the axle. Several bolts in the differential assembly protruded to the rear and could easily puncture the tank. At that point a spark from scraping metal or shredded electrical wiring could readily ignite the spilled

---

[35] Ibid., 101.

[36] Ibid., 101-102.

[37] West's California Reporter, "Grimshaw v. Ford Motor Company," in *The Ford Pinto Case: A Study in Applied Ethics, Business, and Technology*, eds. Douglas Birsch, and John H. Fielder, 253-257 (New York: State University of New York Press, 1994), 253-255.

[38] Mark Dowie, "Pinto Madness," in *The Ford Pinto Case: A Study in Applied Ethics, Business, and Technology*, eds. Douglas Birsch, and John H. Fielder, 15-36 (New York: State University of New York Press, 1994).

[39] Gioia, 104-105.

gasoline.[40] Accounts of pre-production testing revealed that nearly all Pintos subjected to rear end crashes suffered potentially dangerous ruptures of their fuel tanks. Only three of them avoided this failure, and each had been modified to protect the tanks from rupture.[41] The Pinto's notoriety reached its peak in August 1978 when, following a gruesome crash in which three teen-age girls were killed, Ford was charged with reckless homicide in an Indiana court.[42] Though Ford was eventually found not guilty of the charge, the public's confidence in the Pinto and Ford plummeted.

---

[40] Gioia, 100. History echoed itself in disconcerting fashion in the summer of 2000, when Ford Motor Company announced that it's popular and highly profitable Explorer sport utility vehicles were involved in an unusually large number of fatal rollovers. See Matthew L. Wald, "Tread Failures Lead to Recall of 6.5 Million Firestone Tires," *New York Times*, August 10, 2000, https://www.nytimes.com/2000/08/10/business/tread-failures-lead-to-recall-of-6.5-million-firestone-tires.html; Keith Bradsher, "Tire Deaths Are Linked to Rollovers," *New York Times*, August 16, 2000, https://www.nytimes.com/2000/08/16/business/tire-deaths-are-linked-to-rollovers.html. Ford asserted that the Firestone tires used on the vehicles were to blame for the crashes. However, Bridgestone Tire, the owner of Firestone, and independent investigators claimed that the Explorer was at least partly to blame as well. See Joann Muller, and Nicole St. Pierre, "Ford Vs. Firestone: A Corporate Whodunit," *Business Week*, June 11, 2001, 46-47. Sport utility vehicles are heavy and have high centers of gravity. Hence, in the event of tire failure, they are more apt to roll over than passenger cars, and rollovers are often fatal. Critics claimed that this feature of sport utility vehicles is exacerbated by the Explorer's design, which places considerable weight on the left rear tire, since the gasoline tank and the four-wheel drive transfer case are located in the left rear of the vehicle See Keith Bradsher, "Risky Decision/A Special Report; Study of Ford Explorer's Design Reveals a Series of Compromises," *New York Times*, December 7, 2000, https://www.nytimes.com/2000/12/07/business/risky-decision-special-report-study-ford-explorer-s-design-reveals-series.html; Keith Bradsher, "Questions Raised About Ford Explorer's Margin of Safety," *New York Times*, September 16, 2000, https://www.nytimes.com/2000/09/16/business/questions-raised-about-ford-explorer-s-margin-of-safety.html. When the vehicle is traveling at a high rate of speed in hot weather, this tire may fail. As with the Pinto, Ford was eager to rush the Explorer into production in 1989 because it was eager to exploit the growing market for sports utility vehicles. See Bradsher, "Study of Ford Explorer's Design." Also, as was the case with the Pinto, some of Ford's engineers expressed concern over the safety of the vehicle's design. See Bradsher, "Study of Ford Explorer's Design." Though it has never admitted that its product is deficient in any way, Ford quietly began redesigning its Explorers in 1997 to make them safer—and has recently established a sort of driver's education program to demonstrate safe driving techniques to the owners of sport utility vehicles. See Keith Bradsher, "Changes in Ford Explorer Aim at Protecting Other Motorists," *New York Times*, August 4, 2000, https://www.nytimes.com/2000/08/04/business/changes-in-ford-explorer-aim-at-protecting-other-motorists.html; Keith Bradsher, "Explorer Model Raises Doubts About Safety," *New York Times*, April 26, 2001, https://www.nytimes.com/2001/04/26/business/explorer-model-raises-doubts-about-safety.html; Keith Bradsher, "Ford Wants to Send Drivers of Sport Utility Vehicles Back to School," *New York Times*, July 4, 2001, https://www.nytimes.com/2001/07/04/us/ford-wants-to-send-drivers-of-sport-utility-vehicles-back-to-school.html.

[41] Gioia, 100.

[42] Douglas Birsch, "Introduction: The Pinto Controversy," in *The Ford Pinto Case: A Study in Applied Ethics, Business, and Technology*, eds. Douglas Birsch, and John H. Fielder, 3-14 (New York: State University of New York Press, 1994), 5-6.

Though Gioia was Ford's Field Recall Coordinator during a portion of the Pinto's production run, he found nothing in the pattern of consumer complaints or accident reports to reveal a special problem involving the Pinto's fuel tank. However, he was sufficiently moved by the sight of a crashed Pinto's burnt-out hulk that he asked the other members of his department to consider an inquiry. Following a study, the department found no indication of a problem and voted against a recall. Gioia cast his vote with the majority.[43] Gioia points out that this meeting was held before the *Mother Jones* piece appeared or the homicide trial occurred. He also says his office was not aware of the reports of fuel tank ruptures in early testing. He notes in addition that he owned and drove a Pinto during this period and eventually sold it to his sister.[44]

He left Ford in 1975, and for some time thereafter continued to believe that he had made the right decision when he voted against recalling the Pinto. Several years later, he concluded that his decision was mistaken.[45] He spent considerable time and effort reconsidering his actions, and reports that he kept returning to the thought, "Somehow, it seems I should have done *something* different that might have made a difference."[46]

Like Trapp and Stange, Gioia played a small role in a large, complex human organization. As they did, he helped propel a bureaucratic apparatus that sometimes brought harm to human beings. Unlike Trapp and Stange, Gioia's corporate responsibility was to safeguard human life, and he was keenly aware of his burden.[47] Nonetheless, he came to believe that his efforts to preserve human life were insufficient; he did not protect lives that should have been protected. He concluded that he failed because Ford had not instructed him to apply moral principles to the cases he examined. He searched only for patterns which revealed high rates of component failure and a definite causal relation between component failure and incident. He has decided that the Ford could address this deficiency by instructing its Field Recall Coordinators to consider the ethical impact of their decisions.[48]

His proposal reveals that Gioia believes that he was beset by the same problem that undid Trapp, a lack of sufficient moral understanding. Trapp didn't recognize that his actions were morally abhorrent, and Gioia has come to believe he wasn't aware that he needed to employ moral scruples when

[43] Gioia, 103.

[44] Ibid., 104-105, and 107.

[45] Ibid., 106.

[46] Ibid., 105.

[47] Ibid., 102-103.

[48] Ibid., 113-114.

deciding whether to request issuing recalls. His unstated assumption is that he, and others who serve as Field Recall Coordinators, share Trapp's inability to remedy deficiencies of moral understanding. Like Trapp, he applied the standards of judgment his organization had provided, and, like Trapp, he was unable to recognize that he should apply different standards. Gioia presumes, in other words, that he too was a functionary, albeit a functionary charged with protecting human life. He assumes that the functionary's activities can change only when the functionary's role is redefined.

The campus activist, who entered Ford planning to change it, came to believe that Ford should change its job specification. Rather than concluding that Ford needed to seek out more astute and self-assured employees, Gioia decided that Ford needed to alter its procedures in order to introduce the moral sensitivity he found lacking. The Holocaust machine took in men as it found them and transformed them into killers; Gioia appears to believe that corporations must take in human beings as they find them and transform them into employees who are sensitive to the moral demands of their work. He assumes that human institutions can expect their functionaries to become morally accountable only by designing their roles to include it.

There is considerable irony in Gioia's belief that a corporate emphasis on moral standards would have helped him save lives. It is unlikely that enhanced moral sensitivity would have achieved this. A sizeable portion of Gioia's problem was a simple lack of essential information. Gioia did not know of the Pintos' record in preproduction testing. Neither was he aware of concerns a number of engineers voiced about the Pinto's design.[49] If the patterns revealed in consumer complaints and accident reports did not reveal a significant number of failures or a causal linkage for them, it is unclear what grounds Gioia would have possessed for drawing the conclusion that Pintos were unsafe. It is not obvious he would have drawn that conclusion even if he had applied a rigorous standard of concern for human life. Given the information at his disposal, it is difficult to see what basis he had for initiating a recall. Gioia's problem is not equivalent to Trapp's. To fully honor his responsibility to protect human life, Gioia needed more information, not heightened moral sensitivity.

But, there is another issue here, and a crucial one. It is not obvious that Ford was morally irresponsible for designing the Pinto in the way it did, producing millions of copies, then selling them to the public. Unlike the Nazis, Ford did not wish to harm anyone. To the contrary, Ford would likely have been delighted if no one were killed or injured in its vehicles. The factual data are in dispute, but estimates of people killed as the result of igniting Pinto fuel tanks range from 28 to 500. The NHTSA investigation focused on 27

[49] Ibid., 103-104.

deaths.[50] By the end of its production run, there were 3 million Pintos on the road.[51] During the period of its production, several hundred thousand people were killed in automobile accidents in the U.S.[52] People were, and continue to be, killed and injured in every one of the vehicle models Ford produces and in those produced by every auto manufacturer in the world. Automobiles are dangerous devices that kill tens of thousands of people each year in the United States alone and injure many tens of thousands more. Further, small cars are inherently more dangerous than larger ones. Ford's decision to place the gas tank between the rear axle and the rear bumper was a common automotive practice at the time.[53] Finally, as Lee Iacocca, then President of Ford Motor Company, noted pungently, "Safety doesn't sell." [54] More to the point, many people resist taking simple measures, such as wearing seatbelts, driving more slowly and cautiously, or avoiding taking the wheel after drinking. These are all actions that are simple, cost nothing, and would significantly enhance their safety.

The above indicate that it is not obvious that Ford was morally remiss in designing Pintos as it did. Safety is a matter of degree. No car is perfectly safe. Hence, the judgment that a particular model of car is safe or unsafe must be grounded on a variety of considerations that are balanced against one another. Honest and informed people can legitimately disagree on the question of whether Pintos fell below a minimal standard of safety. So, in addition to his lack of information about the Pinto's problems in testing, Gioia lacked clearly defined standards of how safe automobiles must be. Further, since this is a moral problem, not an engineering or manufacturing difficulty, neither Gioia nor Ford Motor Company has authority to address it. These standards can only be devised by representatives of the larger society. Only the nation as a whole can determine how much human life and suffering it is willing to trade for efficiency, style, or comfort.

Hence, despite his belief, Gioia was not beset by the problem that doomed Trapp. Unlike Trapp, he did not need a heightened sense of moral responsibility

---

[50] Office of Defects Investigation Enforcement: National Highway Traffic Safety Administration, "Investigation Report, Phase I: Alleged Fuel Tank and Filler Neck Damage in Rear-End Collision of Subcompact Passenger Cars," in *The Ford Pinto Case: A Study in Applied Ethics, Business, and Technology*, eds. Douglas Birsch, and John H. Fielder, 77-96 (New York: State University of New York Press, 1994), 81.

[51] Birsch, "Introduction: The Pinto Controversy," 6.

[52] The Ford Pinto was manufactured and sold from 1970 to 1980. In that 11-year period, there were 549,447 deaths in traffic accidents in the United States. "Traffic Safety Facts 1997," National Highway Traffic Safety Administration: U.S. Department of Transportation, accessed March 4, 2002, http://www-nrd.nhtsa.dot.gov/pdf/nrd.30/NCSA/TSFAnn/TSF97.pdf.

[53] Birsch, 7-9.

[54] Gioia, 107.

or the recognition that slaughtering innocent people is abhorrent. Gioia needed more information and a clearly defined set of automobile safety standards. Nonetheless, Gioia and Trapp do share something important; both faced moral crises which they as individuals lacked the resources to address successfully.

IV. Moral luck

Some years ago, Bernard Williams and Thomas Nagel introduced the notion of *moral luck*, the idea that the crush of fate and circumstance can turn people into killers, criminals,or...heroes.[55] Certainly Trapp, Stange, and Gioia were buffeted by unsettling twists of moral luck. The Nazi Holocaust apparatus transformed Trapp into a murderer and a criminal. Without it, his moral record would likely have remained unsullied. Stange's services for Captain Novak and Eichmann differed little from those he performed for his entire adult life. Yet, fate transformed his innocuous activity into a critically important element in a machine that brought death to millions. Gioia was sensitive, industrious, and conscientious. He tried to carry out his responsibilities, yet he has come to believe that brute circumstance prevented him from preserving human life as he wished and left him with the sinking feeling that his efforts were not what they should have been. They are all victims of cruel moral luck; had their lives veered in slightly different directions, their moral standing might have shifted as well.

Williams is aware that the idea of moral luck introduces a fundamental shift in our thinking about our moral fates. It links our moral performance tightly to the rest of our lives and shunts aside the conviction that our moral status is immune in some critical way to the forces that shape our lives in other domains.[56] As the league of chastened dotcom workers can attest, success or failure in commercial enterprise depends on blind fate and chance as well as skill, intelligence, and long hours on the job. We recognize that accidents of genetic configuration provide us with short rather than tall stature, a gift for mathematics rather than literature, or physical dexterity rather than hopeless ineptitude. Many believe that accidents of birth play a huge role in delineating our life prospects. Those fortunate to enter wealthy

---

[55] Williams, and Nagel introduced this idea, and ignited a lively controversy on the topic, in papers they delivered at a meeting of the Aristotelian Society. Both have been reprinted. See: Bernard Williams, "Moral Luck," in *Moral Luck*, ed. Daniel Statman, 35-55 (Albany, NY: State University of New York, 1993), and Thomas Nagel, "Moral Luck," in *Moral Luck*, ed. Daniel Statman, 57-71 (Albany, NY: State University of New York, 1993). Statman's useful anthology also includes a number of other illuminating and incisive discussions of this issue.

[56] Williams is well aware of this implication of his view; in fact, he insists on it. See Williams, 35-40.

and successful families are apt to find success in their lives, while those whose lot places them in impoverished environments will have difficulty escaping their surroundings.

Hence, William's notion stands in contrast to Immanuel Kant's influential view that our moral nature and fate are independent of the brute facts which shape our fates in other domains of life.[57] Kant presumed we are always able to avoid becoming murderers. Williams assures us that our moral qualities may sometimes be insufficient to shield us from this fate; in fact, in some circumstances, our moral virtues may propel us to murder. Trapp's character brought him a life of simple decency in Hamburg. In Poland, his character made him a murderer. Trapp is guilty, but guilty through a brutal twist of moral fate. Stange remained the same man he had always been and performed the work he had always performed, but the Nazi regime transformed his skills into vital links in a chain that led to murder. Gioia was conscientious and industrious, but he was unable to preserve human life with the success he desired. Nonetheless, the idea that we only partially control our moral lives is unsettling; it is unsettling to believe that only a few twists and turns of fate stand between us and murder, theft, or callous brutality. We cling to the belief that our moral fates are immune to brute fact, even while we recognize that blind fate is often the key to our financial, physical, or social fortunes. It is entirely true that we sometimes find the moral strength to rise above temptation, self-advantage, or social pressure. But, we are able to rise above our circumstances in other domains of our lives also. And, as happens in life generally, our moral integrity is sometimes crushed by the weight of circumstance. As in other domains of life, we also recognize that some are endowed with exceptional moral strength and integrity. This is why we esteem the likes of Gandhi and Martin Luther King, but we don't expect ordinary people to match their achievements. They serve as ideals whose example we revere and wish to emulate but don't expect to equal.

While the idea of moral luck helps us to comprehend our unease over the fates of Trapp, Stange, and Gioia, it does little to ease our discomfort regarding our moral judgments about them. Though Trapp is a victim of cruel moral luck, he is not less guilty, and our judgment that he is guilty is not less pained. But, that is the way luck works. It is often cruel, heedless, or unjust. *Something went wrong* for Trapp, Stange, and Gioia. Trapp is both decent and guilty. Like a skull cracked by a falling brick, his murderous campaign should not have happened, but it did. A morally better man than Stange would have been aware of the ultimate consequences of his services for Captain Novak, recognize the evil of those consequences, and respond

[57] Immanuel Kant, *Groundwork of the Metaphysic of Morals*, 4: 433, 12ff.

to that evil. In more fortunate circumstances, Gioia would have known more about the Pinto's record and been able to employ a more carefully defined set of automobile safety standards. Trapp, Stange, and Gioia are victims of blind luck.

But, though their lives differ greatly, the bad moral luck that befell each has an element in common: It resulted from the roles they played as workers in sprawling and complex human institutions. In other areas of human life, we do not simply shrug our shoulders and try to get on with our lives when struck with bad luck. We are never entirely helpless, but examine our circumstances and seek ways to remove the causes of our ill fortune. We don't shrug off disease, injury, or economic misadventure. We work to discover the conditions which cause them and then look for ways to eliminate or mitigate the sources of our misfortunes. Vast and enormously powerful human organizations are important for all of us, and they contribute greatly to each individual's burden of moral luck. But, as in other domains of human life, we are bound to consider how we can eliminate or mitigate the bad moral luck they bring. Adam and Eve fell because they were tempted. We are more likely to fall because our lives are entangled with organizations that shape us, shape others, or create troubling moral difficulties for us. What is more, they do so in ways we cannot fully control or understand.

V. Autonomy achieved

At the end of the World War II, the people of West Germany began to face their guilt and sought ways to prevent the crimes of the Nazi era from recurring. They concluded that the ultimate root of Nazi evil was that German society, and the German army in particular, placed enormous emphasis on blind obedience and set no upper boundaries on the fealty which superiors might demand from their subordinates.[58] The issue was addressed squarely in 1955 when, after several false starts and much hesitation, West Germany made the decision, prodded by the Allied Powers, to reestablish its army. The German army had been disbanded at the end of the war, and the need to construct it anew was seen as an opportunity for a fresh start that would avoid the terrible mistakes of the past.[59]

---

[58] Sgt. Michael Maddox, "OSCE Seminar," *SFOR Informer* #93, August 2, 2000, accessed July 12, 2001, http:www.nato.int/sfor/indexinf/93/gertrain/5000730n.html; Donald Abenheim, *Reforging the Iron Cross: The Search for Tradition in the West German Armed Forces* (Princeton: Princeton University Press, 1988), 159.

[59] Count Wolf Baudissin, "The New German Army," *Foreign Affairs* 34, no. 1 (1955): 5-6; Detlef Bald, "Military Reform and the *Innere Führung* in Germany," in *Civil-Military Relations in Post-Communist States: Central and Eastern Europe in Transition*, ed. Anton Bebler, 36-43 (Westport, CT: Praeger, 1997), 38-39.

Innere Führung became a signature doctrine of the reconstituted West German army. It is a term which the Germans insist cannot be translated, but which they nonetheless translate variously as "inner leadership" or "civic and political education," the latter being the preferred rendering.[60] *Innere Führung*, and the array of programs that accompany it, were devised to encapsulate the idea that solders must always remain citizens, with the same rights, responsibilities, and freedoms other citizens possess; they are never simply gears whirling in a vast mechanism.[61] The Germans also understood that their soldiers and citizens required inoculation against racial bias. Hence, the programs also emphasized the dignity and rights of all human beings. In light of their determination to have an army of citizens, the West Germans also made the considered decision to establish a conscript army in which all able-bodied young men would serve. In the army thus devised, soldiers are expected to be politically active and are encouraged to join political parties.[62] They are trained to understand that they should not obey illegal or immoral orders.[63] To inculcate this array of ideas, recruits receive considerable formal training, including sessions in which discussion and disagreement are encouraged.[64]

The Germans also established a training center, the *Zentrum für Innere Führung* in Koblenz. It was given the responsibility of providing elaborate instruction in *Innere Führung* to officers and noncommissioned officers.[65] The educational materials for its endeavors draw on resources provided by civilian academics, clergy, and lawyers.[66] The government also established an independent ombudsman for the military services to which military personnel could send complaints regarding their treatment. Apparently German soldiers keep the ombudsman quite busy, sending a blizzard of tens of thousands of complaints each year.[67] In addition to wielding *Innere Führung* as an instrument to prevent the German army from lapsing into the evil ways of its past, its designers

[60] Maddox; Abenheim, 44 n. 50, and 45.

[61] Donald Hancock, *The Bundeswehr and The National People's Army: A Comparative Study of German Civil-Military Policy* (Denver, CO: University of Denver, 1973), 3.

[62] Baudissin, 9, and 10-12; Bald, 41.

[63] Eric Waldman, *The Goose Step Is Verboten: The German Army Today* (New York: The Free Press of Glencoe, 1964), 144.

[64] Ibid.

[65] Ibid., 140-145.

[66] Ibid., 141.

[67] Bald, 140-141.

hoped the conscripts would return to civilian life with *Innere Führung* firmly in place and spread the idea and its allied practices to the broader society.[68]

This German response mirrors the solution Gioia recommended; the Germans concluded that they could change the beliefs and behavior of individual human beings by devising institutions to shape them in desired ways. They came to believe that individuals sense of moral responsibility and moral standards are shaped by the organizations in which they operate – and they believe that these organizations must reshape themselves and the persons within them in order to avoid recurrence of the tragedies of the Hitler years. Gioia and the founders of the German army appear to share the view that human beings enmeshed in large organizations can be relied upon to be aware of moral standards and view themselves as moral agents only if these organizations explicitly convey these messages to them and the organizations encourage them to act in accordance with their moral scruples. Individual human beings can no longer presume they have full control of their sense of moral responsibility, nor can they presume they are inevitably aware of all fundamental standards of moral conduct. As the Germans discovered, moral accountability and moral standards must be self-consciously nurtured and protected. We can no longer take them for granted; we can no longer presume that our cultural heritage will automatically provide us with a robust sense of moral responsibility, and we cannot longer assume that the moral standards we receive in our youth will prove adequate to the moral problems we face as adults.

Innere Führung, elaborate education in the principle of respect for the lives and dignity of all human beings, and opportunity for independent moral judgment may well have given poor Wilhelm Trapp the personal resources he sorely needed to comprehend the evil of his orders and the strength to refuse them. Though Trapp was in a police battalion, not an army unit, during the World War II, the current German doctrine of *Innere Führung* was devised to address the deficiencies that led to his downfall. However, Trapp's problem is not that he had no moral standards or sense of moral responsibility. He most certainly did. The difficulty is that they were limited by his ingrained deference to authority and lack of imagination. The German doctrine of *Innere Führung* is designed to remedy these limitations.

The West German emphasis on personal responsibility, or something akin to it, is necessary to correct Trapp's failings and those of people like him. However, it does not have resources to address the challenge Stange and Gioia pose. Their difficulties did not arise from failure to appreciate the dignity and inherent worth of the individual, nor did their problems emerge

[68] Baudissin, 13.

from failure to understand that they are citizens and human beings as well as functionaries. *Innere Führung* is focused on the individual and that individual's responsibility for his or her own actions. It is not designed to generate new moral standards when the circumstances of human life spawn moral quandaries. Stange's actions were innocuous; they were not the problem. The problem is that they became part of a chain of human activity that resulted in mass death, but we lack standards which will allow us to assess people's conduct in such cases. Gioia, on the other hand, was keenly aware of his moral responsibility; his difficulty is that he lacked the resources he needed to completely fulfill his moral obligations. He did not have full information about the Pinto's safety record, and he lacked a well-defined set of minimal standards of automobile safety.

The deity has not seen fit to issue new commandments governing the moral responsibilities of individuals who labor in large human organizations, nor has it delivered well-defined standards of automobile safety. Furthermore, the deity has not defined the circumstances under which human beings are obliged to exercise moral agency, and it has not transmitted instructions governing the ways in which human beings should express their agency. As a result, we human beings, as individuals, cultures, or as a species, will have to address these matters using our own resources. Those who direct our vast human organizations want them to function as smoothly and efficiently as possible. This goal is achieved more readily if their employees conduct themselves as cogs and gears. But, people who view themselves as morally accountable will not unthinkingly act as the organization decrees. Neither will they accept the organization's plans and goals without question. People of this sort, people who wish to exercise their moral agency even when they labor as agents of large human institutions, threaten an institution's friction free operation. Left to its own devices, any large organization will work to tune this sensibility out of its workers. In consequence, we may wish to simply bow to this reality and agree that employees bear no moral responsibility for the ultimate consequences of their individual actions. If so, Stange is clearly innocent. He had no obligation to concern himself with the remote consequences of his actions, terrible though they were. On the other hand, we human beings may elect to insist that our moral agency must include responsibility for the ultimate consequences of our acts, even when we play only a small part in bringing them about and they are remote from us. Stone tablets delivered on a mountain peak have no answer to this question. Nor are philosophers of help on this matter. Immanuel Kant would insist that we are obliged to accept responsibility for our actions.[69] True enough. But, Stange's actions were innocuous. The problem is that his actions linked with those

[69] Kant, *Groundwork of the Metaphysic of Morals*, 4: 434-435.

of others to bring evil. Kant doesn't tell us how to resolve that question. We must decide for ourselves how to view Stange and those like him. The decision is ours alone.

In similar fashion, we, as a society, may opt to allow corporations or government agencies to concoct standards of how safe we wish our automobiles, appliances, housing, and streets to be – but we will also have to recognize that allowing this gives them license to weigh our human lives and welfare against cost, efficiency, and style. Or, we may wish to dodge the whole problem by forgoing any effort to establish communal standards in these matters. If so, we will ease away from the demands of moral agency. However, we may also decide that part of the burden of moral agency bequeathed to us by Adam and Eve's theft requires that we accept responsibility for these matters – along with the inconvenience, laborious study, and difficult deliberation this entails. Then we would decide that neither Dennis Gioia nor Ford Motor Company should determine how to strike the balance, for we would have committed ourselves to carrying this burden as a society.

We can no longer take human moral agency for granted. Nor can we continue to presume that individuals have an adequate grasp of the moral standards which should guide their actions. Adam and Eve were accountable for what they did and knew to be wrong. Trapp, Stange, and Gioia didn't understand which moral standards should guide their actions, and it is unlikely that they could be expected to know them. Trapp is guilty for what he did, but Stange and Gioia give pause because their individual actions, innocuous in themselves, were linked to the endeavors of other human beings to bring human suffering. We presently have no clear means of assigning moral guilt or innocence in such cases. However, we cannot overlook these deficiencies, because they ensnarl us all in our day-to-day encounters with vast human institutions.

References

Arendt, Hannah. *Eichmann in Jerusalem: A Report on the Banality of Evil*. New York: Penguin Classics, 1977.

Abenheim, Donald. *Reforging the Iron Cross: The Search for Tradition in the West German Armed Forces*. Princeton: Princeton University Press, 1988.

Bald, Detlef. "Military Reform and the Innere Führung in Germany." in *Civil-Military Relations in Post-Communist States: Central and Eastern Europe in Transition*, edited by Anton Bebler, 36-43. Westport, CT: Praeger, 1997.

Baudissin, Count Wolf. "The New German Army." *Foreign Affairs* 34, no. 1 (1955): 1-13.

Bauman, Zygmunt. *Modernity and the Holocaust*. New York: Cornell University Press, 1989.

Bible, Revised Standard Version. Rock Island, Il: Augustana Book Concern, 1952.

Birsch, Douglas. "Introduction: The Pinto Controversy." In *The Ford Pinto Case: A Study in Applied Ethics, Business, and Technology*, edited by Douglas Birsch, and John H. Fielder, 3-14. New York: State University of New York Press, 1994.

Bradsher, Keith. "Changes in Ford Explorer Aim at Protecting Other Motorists." *New York Times*, August 4, 2000. https://www.nytimes.com/2000/08/04/business/changes-in-ford-explorer-aim-at-protecting-other-motorists.html.

Bradsher, Keith. "Explorer Model Raises Doubts About Safety." *New York Times*, April 26, 2001. https://www.nytimes.com/2001/04/26/business/explorer-model-raises-doubts-about-safety.html.

Bradsher, Keith. "Ford Wants to Send Drivers of Sport Utility Vehicles Back to School." *New York Times*, July 4, 2001. https://www.nytimes.com/2001/07/04/us/ford-wants-to-send-drivers-of-sport-utility-vehicles-back-to-school.html.

Bradsher, Keith. "Risky Decision/A Special Report; Study of Ford Explorer's Design Reveals a Series of Compromises." *New York Times*, December 7, 2000. https://www.nytimes.com/2000/12/07/business/risky-decision-special-report-study-ford-explorer-s-design-reveals-series.html.

Bradsher, Keith. "Questions Raised About Ford Explorer's Margin of Safety." *New York Times*, September 16, 2000. https://www.nytimes.com/2000/09/16/business/questions-raised-about-ford-explorer-s-margin-of-safety.html.

Bradsher, Keith. "Tire Deaths Are Linked to Rollovers." *New York Times*, August 16, 2000. https://www.nytimes.com/2000/08/16/business/tire-deaths-are-linked-to-rollovers.html.

Browning, Christopher R. *Ordinary Men: Reserve Police Battalion 101 and the Final Solution in Poland*. New York: Harper Collins, 1992.

Dowie, Mark. "Pinto Madness." In *The Ford Pinto Case: A Study in Applied Ethics, Business, and Technology*, edited by Douglas Birsch, and John H. Fielder, 15-36. New York: State University of New York Press, 1994.

Goldhagen, Daniel. *Hitler's Willing Executioners*. New York: Knopf, 1996.

Gioia, Dennis A. “Pinto Fires and Personal Ethics: A Script Analysis of Missed Opportunities.” In *The Ford Pinto Case: A Study in Applied Ethics, Business, and Technology*, edited by Douglas Birsch, and John H. Fielder, 97-117. New York: State University of New York Press, 1994.

Hancock, Donald. *The Bundeswehr and The National People’s Army: A Comparative Study of German Civil-Military Policy*. Denver, CO: University of Denver, 1973.

Hilberg, Raul. “German Railroads/Jewish Souls.” *Society* 14, no. 2 (1976): 60-74.

Kant, Immanuel. *Groundwork of the Metaphysic of Morals*. Edited and translated by Mary Gregor, and Jens Timmermann. Cambridge, UK: Cambridge University Press, 2011.

Maddox, Michael. “OSCE Seminar.” *SFOR Informer* #93, August 2, 2000. Accessed July 12, 2001. http:www.nato.int/sfor/indexinf/93/gertrain/5000730n.html.

Muller, Joann, and Nicole St. Pierre. “Ford Vs. Firestone: A Corporate Whodunit.” *Business Week*, June 11, 2001.

Nagel, Thomas. “Moral Luck.” In *Moral Luck*, edited by Daniel Statman, 57-71. Albany, NY: State University of New York, 1993.

National Highway Traffic Safety Administration: U.S. Department of Transportation. “Traffic Safety Facts 1997.” Accessed March 4, 2002. http://www-nrd.nhtsa.dot.gov/pdf/nrd.30/NCSA/TSFAnn/TSF97.pdf.

Office of Defects Investigation Enforcement: National Highway Traffic Safety Administration. “Investigation Report, Phase I: Alleged Fuel Tank and Filler Neck Damage in Rear-End Collision of Subcompact Passenger Cars.” In *The Ford Pinto Case: A Study in Applied Ethics, Business, and Technology*, edited by Douglas Birsch, and John H. Fielder, 77-96.New York: State University of New York Press, 1994.

Sabini, John P., and Mary Silver. “Destroying the Innocent With A Clear Conscience: A Sociopsychology of the Holocaust.” In *Survivors, Victims, and Perpetrators: Essays on the Nazi Holocaust*, edited by Joel Dinsdale, 329-357. Oxfordshire: Taylor & Francis, 1980.

Wald, Matthew L. “Tread Failures Lead to Recall of 6.5 Million Firestone Tires.” *New York Times*, August 10, 2000. https://www.nytimes.com/2000/08/10/business/tread-failures-lead-to-recall-of-6.5-million-firestone-tires.html.

Waldman, Eric. *The Goose Step Is Verboten: The German Army Today*. New York: The Free Press of Glencoe, 1964.

West's California Reporter. "Grimshaw v. Ford Motor Company." In *The Ford Pinto Case: A Study in Applied Ethics, Business, and Technology*, edited by Douglas Birsch, and John H. Fielder, 253-257. New York: State University of New York Press, 1994.

Williams, Bernard. "Moral Luck." In *Moral Luck*, edited by Daniel Statman, 35-55. Albany, NY: State University of New York, 1993.

N. Erinakis · *Conatus* 5, no. 1 (2020): 55-69
DOI: http://dx.doi.org/10.12681/cjp.23180

What Makes Free Will Free: The Impossibility of Predicting Genuine Creativity

Nikos Erinakis
Hellenic Open University, Greece
E-mail address: erinakis.nikos@ac.eap.gr
ORCID ID: http://orcid.org/0000-0001-7188-2993

Abstract

In this paper I argue that Mill's 'Doctrine of Philosophical Necessity' regarding the human will and action cannot apply on all cases, and that the human mind has potentially the capacity to create freely a will or action that, no matter what kind of knowledge we possess, cannot be predicted. More precisely, I argue against Mill's attempt of conjunction between the freedom of the will and the 'Doctrine of Philosophical Necessity' while I attempt a comparison with the relevant Kantian approach. I then claim that a will cannot be free and be predicted at the same time, as both the elements of freedom and unpredictability of the will are founded on the very process of its formulation as an outcome of genuine creativity. I, thus, attempt to propose a more substantially free view of free will and action than the ones presented by the prominent conceptions.

Key-words: *free will; free action; creativity; predictability; compatibilism; philosophical necessity; substantial freedom; John Stuart Mill; Immanuel Kant*

In this paper I explore the deeper relation between the elements of freedom and unpredictability of the human will and action, claiming that for the former to exist, the latter should exist too. Firstly, I deal with the doctrine of 'Philosophical Necessity' as discussed by John Stuart Mill in the case of human will and action and I attempt to prove that certain objections can be raised on whether this doctrine can be applied on all human volitions and actions. Secondly, I suggest that there is a specific kind of human will and action that the doctrine cannot be applied to and I provide a number of relevant examples. I then argue for a view, which aspires to present a more substantially free conception of free will and action than the ones presented by the prominent conceptions, due to the fact that it cannot, by any means, be predicted, since its formulation is an outcome of a process of genuine creativity.

To begin with, I shall concentrate my attention on Mill's chapter "Of Liberty and Necessity" in his work *A System of Logic, Ratiocinative and Inductive* (1843).[1] From the first lines of the chapter, Mill presents the question that he wants to explore and states the two opposing opinions. The question focuses on whether the law of causality applies as strictly to human actions as to other phenomena; that is to say natural phenomena. There are two opinions regarding the possible answer to this question. On the one hand, there is the positive opinion, which is formed by the Necessitarians and suggests that human volition and action are necessary and irrevocable. On the other hand, there is the negative one, which implies that human volition and action is not determined by any causes, but rather it is self-determined. Mill does not hesitate to take a clear position by adopting the affirmative position, that is to say the 'Doctrine of Philosophical Necessity,'[2] of which the main exponent was Joseph Priestley.[3] As a matter of fact, he presents clearly what the 'Doctrine of Necessity' suggests in the case of human volitions and actions[4] and in which way he interprets it. In his own words:

> Correctly conceived, the doctrine called Philosophical Necessity is simply this: that, given the motives which are present to an individual's mind, and given likewise the character and disposition of the individual, the manner in which he will act might be

---

[1] John Stuart Mill, *A System of Logic: Ratiocinative and Inductive*, in *The Collected Works of John Stuart Mill*, ed. J. M. Robson (Toronto: Toronto University Press, 1974) [hereinafter Mill, *System of Logic*].

[2] Mill, as well as Herbert Spencer, was greatly influenced by the form of determinism expressed by the view of 'Philosophical Necessity.' The main advocator of 'Philosophical Necessity' was Joseph Priestley; see Joseph Priestley, *The Doctrine of Philosophical Necessity Illustrated; Being an Appendix to the Disquisitions Relating to Matter and Spirit. To Which is Added an Answer to the Letters on Materialism, and on Hartley's Theory of the Mind* (London: Printed for J. Johnson, 1777). African Spir believed that after Priestley's work, it was pointless to discuss the issue of free will further; see African Spir, *Denken und Wirklichkeit. Versuch einer Erneuerung der kritischen Philosophie* (Leipzig: J. G. Findel, 1873), 162.

[3] For more information on Priestley's 'Doctrine of Philosophical Necessity' see Clarke Garrett, "Joseph Priestley, the Millennium, and the French Revolution," *Journal of the History of Ideas* 34, no. 1 (1973): 51-66; also Isaac Kramnick, "Eighteenth-Century Science and Radical Social Theory: The Case of Joseph Priestley's Scientific Liberalism," *Journal of British Studies* 25, no. 1 (1986): 1-30; Robert E. Schofield, *The Enlightened Joseph Priestley: A Study of His Life and Work from 1773 to 1804* (University Park: Pennsylvania State University Press, 2004).

[4] Arthur Schopenhauer wrote that "[...] no writer has presented the necessity of acts of will so thoroughly and convincingly as Priestley [...] If anyone is not convinced by this supremely clearly and accessibly written book, his understanding must really be paralysed by prejudices," and that the work contributed to Kant who considered "the complete necessity of acts of will as a settled matter to which no further doubt could pertain." Arthur Schopenhauer, *On the Freedom of the Will* (Trondheim: Royal Norwegian Society of Sciences, 1839), respectively 77, 81.

> unerringly inferred; that if we knew the person thoroughly, and knew all the inducements which are acting upon him, we could foretell his conduct with as much certainty as we can predict any physical event. This proposition I take to be a mere interpretation of universal experience, a statement in words of what every one is internally convinced of.[5]

Although, this seems to be a very well structured and fortified statement, as it seems to be referring to all possible factors that can be influencing human volition, we cannot ignore the fact that it seems very general and vague and thus a number of objections can be raised against it. Mill concentrates on three critical internal factors of the agent: the motives, the character and the disposition, and he summarizes all the external factors to one: all the inducements that can possibly act upon the agent. He implies that if potentially we knew all these factors, they would be enough to predict exactly the volition or the action of the agent. In fact, this statement seems vague since Mill does not make clear what other exact elements he considers to be included in the motives, character and disposition of the agent. [6]

The first question that can be directly raised is how can we actually attain this kind of knowledge. Mill defends his position by making clear that he does not support that we can actually possess this knowledge but rather he supports that if we could possibly have it, we could definitely produce such predictions successfully. That is, given the knowledge of the relevant preconditions and laws, it is necessary that a person acts in a certain way, and a fully informed observer could predict precisely this person's will and action. In other words, ceteris paribus, what a fully informed observer predicts about the person's will and action, cannot but necessarily happen.

Even then, however, if we are to accept this hypothesis, another, even more critical, question can be raised: Is this knowledge adequate? Are these factors enough to predict human volition and action? Even if we had the ability to approach this kind of knowledge, that is to say to know all

---

[5] Mill, *System of Logic*, 836.

[6] In regard to motives, which play such a crucial role on the determination of will in Mill consider Norman Wilde's argument referring to Kant's notion of *Good Will*: "Motives, so far as they are internal, a supposed free will or autonomous self, cannot be subjected to observation and experiment, and hence must be ignored in considering the principles of moral judgment. The name of Kant, on the other hand, has always been connected with a system the exact opposite of the one just described. The celebrated opening of the Grundlegung has always been accepted as the keynote of his moral theory. 'Nothing can possibly be conceived in the world, or even out of it, which can be called good without qualification, except a Good Will.' Here we have the direct contrary of the utilitarian position that the will has no value save in relation to the effects produced by it." Norman Wilde, "Kant's Relation to Utilitarianism," *The Philosophical Review* 3, no. 3 (1894): 292.

these factors, would it be enough for us to predict exactly what a person would will and how she would act?

In fact, in this paper I argue that if we were to know the motives, the character, the disposition and all the possible inducements that can act upon a person, we might be able to predict all kinds of her volitions and actions, except one, which we can definitely never predict: the will and/or act that nobody ever before has imagined, thought, formulated and/or done. This said, although Mill's definition seems complete, he does not take into account the element of *genuine creativity*. By the element of and/or capacity for genuine creativity I refer to the *creative process* of a will or action of a person that occurs for the first time, in the sense that it is, even in the slightest degree, *historically* and *psychologically novel, unique* and *original*.

Mill implies that the aforementioned factors, i.e. the motives, the character, the disposition and all the possible inducements, are enough to predict all kind of volitions and actions of a person. Mill's aspiration to provide a solution to the problem of compatibility between determinism and free will is based on what he refers to as a misleading association to the notion of *necessity*. Here we need to clarify the difference between the idea that actions occur necessarily and the idea that actions are predetermined and agents have no influence on them. Given this distinction, the doctrine of necessity, which for Mill is compatible with determinism,[7] is differentiated from the doctrine of fatalism,[8] which is not compatible with determinism. However, even by taking into account these remarks, the crucial question remains: How can one predict a will or an action that has never occurred before? How can you predict something that you cannot even imagine? Thus, Mill's aforementioned factors seem to be inadequate, that is to say they are neither sufficient nor necessary for predicting human will and action, because they cannot predict the spontaneous outcomes of genuinely creative processes when they occur for the first time both in one's personal and in collective history. Nevertheless, after the first time that such a volition or action takes place, it is added in the map of possible options of choices and the doctrine of Philosophical Necessity may possibly apply upon it too.

---

[7] Notice that Mill's solution to the Problem of Necessity (i.e. the problem of determinism and freedom of the will), is one of the first compatibilist conceptions of free will. For more on this, see Elijah Millgram, "John Stuart Mill, Determinism, and the Problem of Induction," *Australasian Journal of Philosophy* 87, no. 2 (2009): 183-199; George Boolos, "On Second-Order Logic," *The Journal of Philosophy* 72, no. 16 (1975): 509-527; Janice Carlisle, *John Stuart Mill and the Writing of Character* (Athens, GA: University of Georgia Press, 1991).

[8] For more information on Fatalism, see Alfred Jules Ayer, "Fatalism," in *The Concept of a Person and Other Essays* (London: McMillan, 1963), 235-268; also Richard Taylor, "Fatalism," *The Philosophical Review* 71, no. 1 (1962): 56-66.

Perhaps we could understand better this claim if we considered an example from art. Let us imagine the first painter who ever drew truly abstractly, e.g. one of the pioneers of abstract expressionism. Making the hypothesis that people of her/his time knew completely everything about her/his life, motives and disposition and all the inducements acting upon her/him, they could have predicted what kind of paintings she/he would have intended to draw and in which way, but they could never have predicted that she/he would drew abstractly, outside of any already known forms and techniques. Let us here underline the fact that this kind of genuine creativity can be noticed not only in individual persons but also in collective movements etc.

It might also be useful to think of the case of Medea.[9] Let us assume, in a rather simplified way, that firstly, Medea's context was a historically existing one, and secondly, that no mother before Medea had ever reached the state of killing her children in order to revenge her husband. One could predict that when a wife feels cheated, she might want to take revenge and this revenge might be extreme, but before Medea no person could assume that a wife could possibly murder her children for revenge. In other words, I believe that there are many cases of different kinds of human volitions and actions (from aesthetically pleasing to morally repulsive), which even if we possess the knowledge of all relevant factors, both internal and external, which influence them, we are not able to predict them (e.g. something that happens for the first time in the known history of humankind and is, in such a way or degree, unique that no other person could have acknowledgingly ever imagined it).

Moreover, one cannot but wonder if Mill took also into account cases of volitions and actions of mentally disordered persons. In cases of severe mental disorders nobody can predict what the person will or is able to do, even if we possessed all the knowledge that Mill suggests that is needed to do so. More precisely, the motives, the character and the disposition of a person together with all the inducements that can act upon her are not enough in order to predict the volition or an act of a seriously disordered mind. To begin with, a disordered mind does not work through a coherent causal system that we may be able to identify, like it has been mainly suggested for the human mind in general (although there are a number of lines of thought that argue against this supposition). But even if we supported that a disordered mind does work based on coherent causal laws too, we should accept that it does so based on much less obvious causal systems than other minds do. The key point here is that a disordered mind may have the ability to operate through novel ways of thinking and acting, and, while seeming irrational to other minds, to have the capacity for an alternative form of genuine creativity. Therefore, despite the

[9] Medea was the wife of Jason and mother of Mermeros and Pheres. When Jason betrayed her (he left her for Glauce, Creon's daughter), she killed her children in order to revenge him.

amount of knowledge that one may possess for mentally disordered persons, one remains unable to predict the outcomes of at least a certain number of them that retain, at least, an alternative capacity for genuine creativity.

Nevertheless, a Necessitarian could argue that by expanding these factors or deepen our knowledge on them, we could be able at some point to predict all cases of both mentally healthy and mentally disordered minds. But then we are once again faced with the paradox stated above. How can you make a prediction[10] about something that you do not know yet that exists, i.e. it is not a part of any known conceptual space[11] (such as the case of a genuine creative volition or act formulated through a creative process of an either mentally healthy or a mentally disordered person)? How can you predict something that you have never acknowledged, experienced or noticed before and never had a clue about it? If you cannot even imagine something, then certainly you cannot predict it. Could any infectious disease specialist ever predict or speak of a potentially pandemic virus that has never occured before? Could she recognize it or even conceive its existence?

Therefore, in the case of either a healthy mind or a disordered mind that can give birth to volitions and actions based on factors which other human minds cannot even imagine, nothing can be predicted. However, let us keep in mind that such volitions and acts cannot be predicted solely for the first time that they occur. Their key element is the combination of originality, uniqueness and singularity of their nature. They manage to operate in a way that no mind has ever acknowledgingly operated before. They create a novel way of willing and a novel way of acting. Even if neuroscience, through its astonishingly promising paths, manages to read the human brain in its whole and understand in total the ways in which it operates, there will always exist a potentiality of a human being with a certain – biological or not – brain 'distortion,' that will lead to the formation of a volition or an action that could not by any means be predicted. This kind of distortion is one of the cases that confound the element of prediction as Mill and the Necessitarians conceive it, i.e. a conception in which free will is not substantially free, since which capacity or attitude can be truly free when it can be predicted based on the very factors that formulated it. Spontaneity is a substantial element of freedom, and one that requires some form of creativity. Genuine creativity,

---

[10] For a thorough argumentation on the subject of *predictability* and its relation to *determinism*, see Daniel J. O'Connor, "Determinism and Predictability," *British Journal for the Philosophy of Science* 7, no. 28 (1957): 310-315; Peter Herbst, "Freedom and Prediction," *Mind* 66, no. 261 (1957): 1-27; Haskell Fain, "Prediction and Constraint," *Mind* 67, no. 267 (1958): 366-378.

[11] For further clarification on the concept of the *conceptual space*, see Margaret Boden, *The Creative Mind: Myths and Mechanisms* (London: Routledge, 2004); also *Dimensions of Creativity*, ed. Margert Boden (Cambridge, MA: MIT Press, 1994).

thus, especially because its outcomes cannot be predicted, is the key element for both the existence and unpredictability of freedom of the will.

When Mill argues that, "We are exactly as capable of making our own character, *if we will*, as others are of making it for us,"[12] Mill clearly believes that if we desire to change ourselves then we have the ability to do so, i.e. experience shows that we can affect our habits and attitudes and that this desire resides in our selves.[13] Leaving aside for now, how deeply problematic may be to speak of an existing self[14] as a robust, coherent entity,[15] the aforementioned view of Mill operates as the substance of his doctrine of free will:

> [...] that what is really inspiring and ennobling in the doctrine of free will, is the conviction that we have real power over the formation of our own character; that our will, by influencing some of our circumstances, can modify our future habits or capabilities of willing. All this was entirely consistent with the doctrine of circumstances, or rather, was that doctrine itself, properly understood.[16]

For Mill, therefore, we are free in the sense that we can become those who we want to be, although any fully informed bystander not only could know *what we would desire to become*, but also, more importantly, *what*

---

[12] John Stuart Mill, *The Collected Works of John Stuart Mill*, ed. John M. Robson (Toronto: University of Toronto Press, 1963-1991), vol. 8, 841 [hereinafter Mill, *Collected Works*].

[13] Ibid.

[14] For more thoughts on Mill's idea of the self, see Wendy Donner, *The Liberal Self. John Stuart Mill's Moral and Political Philosophy* (Ithaca & London: Cornell University Press, 1991); Fred R. Berger, *Happiness, Justice, and Freedom: The Moral and Political Philosophy of John Stuart Mill* (Berkeley & Los Angeles: University of California Press, 1984); John Skorupski, *Why Read Mill Today?* (London & New York: Routledge, 2006); James O. Urmson, "The Interpretation of the Moral Philosophy of J. S. Mill," *The Philosophical Quarterly* 3, no. 10 (1953): 33-39; Samuel Clark, "Love, Poetry, and the Good Life: Mill's *Autobiography* and Perfectionist Ethics," *Inquiry* 53, no. 6 (2010): 565-578.

[15] According to another line of thought, not without several problems itself too, which has been espoused, from different directions, by empiricists, neuroscientists, feminists and post-modern thinkers, the self is constantly changing and incoherent and, in general, can be considered an illusion. For more, see Daniel C. Dennett, "The Self as a Center of Narrative Gravity," in *Self and Consciousness: Multiple Perspectives*, eds. F. Kessel, P. Cole, and D. Johnson, 103-115 (Hillsdale, NJ: Erlbaum, 1992); Daniel C. Dennett, Maxwell Bennett, Peter Hacker, and John Searle, *Neuroscience and Philosophy: Brain, Mind, and Language* (New York: Columbia University Press, 2007); Marilyn A. Friedman, "Autonomy and the Split-Level Self," *The Southern Journal of Philosophy* 24, no. 1 (1986): 19-35; Bruce Hood, *The Self Illusion: Why There Is no 'You' Inside Your Head* (London: Constable, 2012); Diana T. Meyers (ed.), *Feminists Rethink the Self* (Boulder: Westview Press, 1997).

[16] Mill, *Collected Works*, 1, 177.

we are going to become. He claims that: "I dispute therefore altogether that we are conscious of being able to act in opposition to the strongest present desire or aversion."[17] Hence, since one cannot counteract one's much *effective desire*, one cannot but be necessarily determined by it, and thus the outcome of this determination can be fully and accurately predicted. In a comparison with Kant's thought,[18] Norman Wilde claims that:

> If the desire for happiness were to be admitted as the principle for the determination of the will, no necessary law could be formulated; since pleasurable consciousness is the result of experience, from no examination of which could any necessary law be derived. In other words, the principle would be Heteronomy instead of Autonomy, which alone gives a basis for distinctly moral judgment. To allow a material principle such as pleasure to determine the will would be to place the Ego under natural laws in which no freedom is possible – and if no freedom, then for Kant, no morality is possible, since "the ratio essendi of the moral law is freedom," i.e., self legislation.[19]

Consequently, Mill asserts that the fact that others can predict the volition or an action of an agent does not mean that the agent does not have free will;

---

[17] Mill, *Collected Works*, 9, 453.

[18] However, it is important to keep in mind here that, in general, Mill looks into morality as a social practice and not as autonomous self-determination by reason, like Kant. For further argumentation on this, see Rocer Hancock, "Ethics and History in Kant and Mill," *Ethics* 68 (1957): 56-60; Roderick Ninian Smart, "Negative Utilitarianism," *Mind* 67, no. 268 (1958): 542-543; John Jamieson Carswell Smart, and Bernard Williams, *An Outline of a System of Utilitarian Ethics* (Cambridge*:* Cambridge, 1973); Jan Narveson, *Morality and Utility* (Baltimore, MD: Johns Hopkins Press, 1967).

[19] Norman Wilde, "Kant's Relation to Utilitarianism," *The Philosophical Review* 3, no. 3 (1894): 289-304, 292-293. In addition, of interesting relevance is Dennis A. Rohatyn's claim that: "In view of these doctrinal affiliations, it is surprising to find Mill attacking Kant's conception of free will and moral autonomy elsewhere (see [5], Ch. 26, esp. 298-299). Here Mill suggests that the Kantian conception is at odds with Mill's own ideas on the formation of character, or 'Ethology,' a topic which Mill adumbrated at various times, and discussed embryonically in his System of Logic (Bk. VI, Ch. v, para. 4). Mill's own resolution of the free-will question in the face of his idealized social science of man is itself muddled, hinging as it does on a vexed distinction between external (physical) and internal (agent's own volition) causality. But this confusion surfaces in Kant's third antinomy, as well, where it is made to accord with the idea of the will legislating maxims for itself (see [4], A444/B472, and also [3], 421). So even Mill's criticisms of Kant turn out to be based on theories with which Kant would agree, or toward which the latter was moving." Dennis A. Rohatyn, "Mill, Kant, and Negative Utility," *Philosophia* 5 (1975): 517.

contrariwise the agent remains free to will whatever she prefers.[20] What I am trying to highlight here is that, in my opinion, the notion of *genuine creativity* to which I am referring, is a capacity of the human mind that *operates before*, and in a deeper level than, the *formulation of one's desires*. The very source of the elements of freedom and unpredictability of the will resides in the *capacity for creative processes*, which in turn can create creative desires, and not in plain desires of the person as it has been suggested.

Besides this, however, if other people can possibly predict, and thus know, in which ways one is going to act, then potentially one is able to know it too (either by being provided with this knowledge by others, or by using the same factors that the others used to make the prediction). Hence, if you know exactly how you are going to act – even if you can act in whichever way you want – you are not free, but constrained by the factors that are supposed to form your will. In addition to this, you also lose the freedom of changing your mind at the last moment. Thus, we are faced with a paradox here. Whilst Mill supports the existence of free will, he supports also the doctrine of the Philosophical Necessity on human actions, which based on the above arguments restricts the free will of the agent. However, those who desire to object this view may argue that even a last moment change of mind can be predicted if we know all the previously mentioned factors. But is it not restrictive to accept the fact that individuals are capable of thinking and acting only in the ways and on the options, which these specific factors can give birth to (even if these ways and options are abundant) and, thus, to exclude the possibility of not only *creating a genuine will or action* but also *creating a genuine way and/or option of willing and acting* to begin with?

Furthermore, if for a moment we assume that the agent may know all the factors that influence her will or action, then she can possibly change her will, after the prediction is formulated, in a way that cannot be predicted. If we take this idea one step further, we can support that not only an agent can alter her will, but also she can alter the factors that influence her volition and her actions. For instance, if an individual knows the factors that are used in order for her actions to be predicted according to Mill, then she can possibly alter them in way which resides out of the possible predictions map. In other words, she can barricade her will from the factors that influence it and base her actions on innovative factors. This might have as a result the cancelation of the validity of the possible predictions.

Aristotle wrote, "Nor is there any definite cause for an accident, but only chance (τυχόν), namely an indefinite (ἀ όριστον, aoriston) cause."[21]

---

[20] Mill, *Collected Works*, 9, 548.

[21] Aristotle, *Metaphysics*: *Book V,* trans. Hugh Tredennick, and G. Cyril Armstrong (Cambridge,

Following from this, we may conceive the volitions or actions which occur for the first time as nor outcomes of causal laws neither of chance, but rather as an outcome of an *aoriston* (indefinite/indetermined) cause, which immediately after they occur become *orismenoi* (definite/determined).[22] Given this, I do not intend to suggest that the element of chance or luck is the reason why not all human actions can be predicted. The outcomes of genuine creativity should not, by any means, be understood as occurring from an element of chance or luck alone. Besides, we may here consider the concept of *moral luck*[23] as proposed by Bernard Williams[24] and further developed by Thomas Nagel[25] in their renowned pair of articles, based on Kant's idea of the Good Will in the *Groundwork for the Metaphysics of Morals*,[26] claiming that our everyday judgments and practices commit us to the existence of it. However, a number of thinkers in the free will debate have argued against the existence of causal moral luck, while proposing a distinctive metaphysical account of human agency.[27] This view is known as the "Agent-Causal Libertarianism" and the core idea is that the actions, or at least the formation of intentions, are caused by the persons themselves, without the persons being caused to do so.

MA: Harvard University Press, 1935), 1025a25.

[22] For thoughts on Mill's relation to Aristotle on this topic, see Martha C. Nussbaum, "Mill between Aristotle & Bentham," *Daedalus* 133, no. 2 (2004): 60-68.

[23] Thomas Nagel defines *moral luck* as follows: "Where a significant aspect of what someone does depends on factors beyond his control, yet we continue to treat him in that respect as an object of moral judgment, it can be called moral luck." Thomas Nagel, Mortal Questions (New York: Cambridge University Press, 1979), 59.

[24] Bernard Williams, *Moral Luck* (Cambridge: Cambridge University Press, 1981); Bernard Williams, "Postscript," in *Moral Luck*, ed. D. Statman, 251-258 (Albany: State University of New York Press, 1993).

[25] Thomas Nagel, *Mortal Questions* (New York: Cambridge University Press, 1979); Thomas Nagel, *The View from Nowhere* (New York: Oxford University Press, 1986).

[26] Immunity from luck has been thought by many to be part of the very essence of morality and it has found its inspiration in Kant: "A good will is not good because of what it effects or accomplishes, because of its fitness to attain some proposed end, but only because of its volition, that is, it is good in itself [...] Even if, by a special disfavor of fortune or by the niggardly provision of a step motherly nature, this will should wholly lack the capacity to carry out its purpose – if with its greatest efforts it should yet achieve nothing and only the good will were left (not, of course, as a mere wish but as the summoning of all means insofar as they are in our control) – then, like a jewel, it would still shine by itself, as something that has its full worth in itself. Usefulness or fruitlessness can neither add anything to this worth nor take anything away from it." Immanuel Kant, *Groundwork for the Metaphysics of Morals*, ed. and trans. Mary J. Gregor (Cambridge: Cambridge University Press, 1996), 4: 394.

[27] See Roderick Chisolm, "Human Freedom and the Self," *The Lindley Lecture*, University of Kansas (1964); Randolph Clarke, "Toward a Credible Agent-Causal Account of Free Will," *Noûs* 27 (1993): 191-203; Richard Taylor, *Action and Purpose* (Englewood Cliffs, NJ: Prentice Hall, 1966); Timothy O'Connor, *Persons and Causes: The Metaphysics of Free Will* (New York: Oxford University Press, 2000).

Thus, an agent, while exercising her own causal powers, is an undetermined cause of her own intentions. This said, in my opinion, the *source of one's causal powers* that ensures the *freedom* and *unpredictability* of one's *intentions*, *volitions* and *actions*, by rendering one *the undetermined cause* of them, is the capacity for *genuine creativity*.

Following from the above, I am of the opinion that a will or act which take place for the first time may occur both through intentional rational reflection and/or through imaginativeness[28] of the agent and still without being able to be predicted, since the agent, at least in cases of genuine creativity, operates for sure as the undetermined cause of them. On the relationship of rationality, freedom and causal relationship let us think here Kant's main argument in the *GMS* that regarding oneself as rational implies regarding oneself as free.[29] However, whether through Kant's thought we can succeed in showing that firstly, regarding oneself as free is incompatible with accepting universal causal determinism and secondly, regarding oneself as rational is incompatible with accepting universal causal determinism, remain open questions for further investigation.[30]

This said, I accept that the doctrine of Philosophical Necessity might be true for a majority of human volitions and actions except those of authentic, genuine creativity, which, by definition, occur for the first time, i.e. they are characterized

---

[28] For a historical review and discussion of the notion of imaginativeness, see James Eric Grant, *The Critical Imagination* (Oxford: Oxford University Press, 2013).

[29] For thorough argumentation on this, see Corina Mieth, and Jacob Rosenthal, "'Freedom must be Presupposed as a Property of the Will of All Rational Beings' (GMS III, 2)," in Immanuel Kant, *Groundwork for the Metaphysics of Morals*, eds. Cristoph Horn, and Deiter Schoenecker, 247-284 (Berlin and New York: de Gruyter, 2006); Richard Mervyn Hare, *Freedom and Reason* (Oxford: Oxford University Press, 1963); Allen Wood, *Kantian Ethics* (Cambridge: Cambridge University Press, 2007); Alvin I. Goldman, "The Compatibility of Mechanism and Purpose," *The Philosophical Review* 78, no. 4 (1969): 468-482; William Hasker, "The Transcendental Refutation of Determinism," *The Southern Journal of Philosophy* 11, no. 3 (1973): 175-183.

[30] For now, however, Wilde's suggestion operates in an illuminating manner: "To say that man is to act from a principle which he can will to be universal law – to make him the supreme law-giver to himself, is to place him above the reach of all law other than his own, or that which he freely makes his own. If, then, as a rational being he cannot be subject to anything without him, he is an end unto himself – an absolute end so far as rational. All men, moreover, as partakers of a common reason, must also be regarded as ends, and so treated, i.e., the ends of each individual must be made the ends of the particular subject. Thus from the idea of law Kant passes to its equivalent, the idea of the self, since the form of law is but the expression of the nature of the rational self. The term 'law' in this connection rather confuses the real identity of these two notions, since it is generally used in Kantian ethics to denote a command. But it is thus used in relation to the lower self as constrained, and not of the higher or noumenal self as the source of law. In regard to this latter, it denotes merely natural law, habit, or nature. The moral law, as issued to the self as object in the form of command, is really the expression of true nature of the rational self as subject. It is only as imposed upon the self regarded as inclination, that it takes the form of command." Norman Wilde, "Kant's Relation to Utilitarianism," *The Philosophical Review* 3, no. 3 (1894): 296.

by historical and psychological novelty. The element of indeterminate surprise of genuine creativity here is more than crucial. This direction places the core argument of this paper on the thin line between compatibilism[31] and libertarianism. More specifically, I certainly support the existence of free will and action (as, in general, Mill does too), but I disagree on the fact that it can always be predicted, and thus I argue for the existence of a more substantially free form of volition. I support the idea that incidents of reality can be considered as determined (orismenoi), but only after these incidents have occurred before at least once. More precisely, a will or action that is formulated through one's capacity for genuine creativity is considered indetermined (aoriston) until it takes place for the first time, when it immediately becomes determined (orismeni) for all the following times that will occur.

In this paper I attempted to suggest that there are cases of human will and action that cannot be predicted even if we possessed all the knowledge that Mill and others believed to be adequate, i.e. both sufficient and necessary, in order to do so. An exemplified case on which this paper was based is the formation of volitions or actions, which occur for the first time ever, through the capacity for genuine creativity. I maintained that in this case, Mill's doctrine of Philosophical Necessity cannot apply, and that the human mind, through its creative processes, has potentially the capacity to formulate a free will or action that, no matter what kind of knowledge we possess, cannot be predicted. I argued, thus, against Mill's attempt of conjunction of the freedom of the will and the doctrine of Philosophical Necessity, while I claimed that a will cannot be free and be predicted at the same time, as both the elements of freedom and unpredictability of the will are founded on the very *process* of its formulation *as an outcome* of *genuine creativity*.

References

Aristotle. *Metaphysics*: *Book V*. Translated by Hugh Tredennick, and G. Cyril Armstrong. Cambridge, MA: Harvard University Press, 1935.

Ayer, Alfred Jules. "Fatalism." In *The Concept of a Person and Other Essays*. London: Macmillan, 1963.

Berger, Fred R. *Happiness, Justice, and Freedom: The Moral and Political Philosophy of John Stuart Mill*. Berkeley & Los Angeles: University of California Press, 1984.

[31] For further discussion on compatibilism, i.e. the reconcilability of determinism and free will, see John V. Canfield, "The Compatibility of Free Will and Determinism," *The Philosophical Review* 71, no. 3 (1962): 352-368; Harald Ofstad, *An Inquiry into the Freedom of Decision* (London: George Allen & Unwin, 1961).

Boden, Margaret, ed. *Dimensions of Creativity*. Cambridge, MA: MIT Press, 1994.

Boden, Margaret. *The Creative Mind: Myths and Mechanisms*. London: Routledge, 2004.

Boolos, George. "On Second-Order Logic." *The Journal of Philosophy* 72, no. 16 (1975): 509-527.

Canfield, John V. "The Compatibility of Free Will and Determinism." *The Philosophical Review* 71, no. 3 (1962): 352-368.

Carlisle, Janice. *John Stuart Mill and the Writing of Character*. Athens, GA: University of Georgia Press, 1991.

Clarke, Randolph. "Toward a Credible Agent-Causal Account of Free Will." *Noûs* 27, no. 2 (1993): 191-203.

Clark, Samuel. "Love, Poetry, and the Good Life: Mill's *Autobiography* and Perfectionist Ethics." *Inquiry* 53, no. 6 (2010): 565-578.

Chisolm, Roderick. "Human Freedom and the Self." In *The Lindley Lecture*. University of Kansas, 1964.

Dennett, Daniel C. "The Self as a Center of Narrative Gravity." In *Self and Consciousness: Multiple Perspectives,* edited by F. Kessel, P. Cole, and D. Johnson, 103-115. Hillsdale, NJ: Erlbaum, 1992.

Dennett, Daniel C., Maxwell Bennett, Peter Hacker, and John Searle. *Neuroscience and Philosophy: Brain, Mind, and Language*. New York: Columbia University Press, 2007.

Donner, Wendy. *The Liberal Self. John Stuart Mill's Moral and Political Philosophy*. Ithaca & London: Cornell University Press, 1991.

Fain, Haskell. "Prediction and Constraint." *Mind* 67, no. 267 (1958): 366-378.

Friedman, Marilyn A. "Autonomy and the Split-Level Self." *The Southern Journal of Philosophy* 24, no. 1 (1986): 19-35.

Garrett, Clarke. "Joseph Priestley, the Millennium, and the French Revolution." *Journal of the History of Ideas* 34, no. 1 (1973): 51-66.

Grant, James Eric. *The Critical Imagination*. Oxford: Oxford University Press, 2013.

Goldman, Alvin I. "The Compatibility of Mechanism and Purpose." *The Philosophical Review* 78, no. 4 (1969): 468-482.

Hancock, Rocer. "Ethics and History in Kant and Mill." *Ethics* 68 (1957): 56-60.

Hare, Richard Mervyn. *Freedom and Reason*. Oxford: Oxford University Press, 1963.

Hasker, William. "The Transcendental Refutation of Determinism." *The Southern Journal of Philosophy* 11, no. 3 (1973): 175-183.

Herbst, Peter. "Freedom and Prediction." *Mind* 66, no. 261 (1957): 1-27.

Hood, Bruce. *The Self Illusion: Why There Is no 'You' Inside Your Head*. London: Constable, 2012.

Kant, Immanuel. *Groundwork for the Metaphysics of Morals*. Edited and translated by Mary J. Gregor. Cambridge: Cambridge University Press, 1996.

Kramnick, Isaac. "Eighteenth-Century Science and Radical Social Theory: The Case of Joseph Priestley's Scientific Liberalism." *Journal of British Studies* 25, no. 1 (1986): 1-30.

Meyers, Diana T., ed. *Feminists Rethink the Self*. Boulder: Westview Press, 1997.

Mieth, Corinna, and Jacob Rosenthal. "'Freedom must be Presupposed as a Property of the Will of All Rational Beings' (GMS III, 2)." In Immanuel Kant, *Groundwork for the Metaphysics of Morals*, edited by Cristoph Horn, and Deiter Schoenecker, 247-284. Berlin & New York: de Gruyter, 2006.

Mill, John Stuart. *The Collected Works of John Stuart Mill*. Edited by John M. Robson. Toronto: University of Toronto Press, 1963-1991.

Mill, John Stuart. *A System of Logic: Ratiocinative and Inductive*. In *The Collected Works of John Stuart Mill*, edited by J. M. Robson. Toronto: Toronto University Press, 1974.

Millgram, Elijah. "John Stuart Mill, Determinism, and the Problem of Induction." *Australasian Journal of Philosophy* 87, no. 2 (2009): 183-199.

Nagel, Thomas. *Mortal Questions*. New York: Cambridge University Press, 1979.

Nagel, Thomas. *The View from Nowhere*. New York: Oxford University Press, 1986.

Narveson, Jan. *Morality and Utility*. Baltimore, MD: Johns Hopkins Press, 1967.

Nussbaum, Martha C. "Mill between Aristotle & Bentham." *Daedalus* 133, no. 2 (2004): 60-68.

O'Connor, Daniel J. "Determinism and Predictability." *British Journal for the Philosophy of Science* 7, no. 28 (1957): 310-315.

O'Connor, Timothy. *Persons and Causes: The Metaphysics of Free Will*. New York: Oxford University Press, 2000.

Ofstad, Harald. *An Inquiry into the Freedom of Decision*. London: George Allen & Unwin, 1961.

Priestley, Joseph. *The Doctrine of Philosophical Necessity Illustrated; Being an Appendix to the Disquisitions Relating to Matter and Spirit. To Which is Added an Answer to the Letters on Materialism, and on Hartley's Theory of the Mind*. London: Printed for J. Johnson, 1777.

Rohatyn, Dennis A. "Mill, Kant, and Negative Utility." *Philosophia* 5 (1975): 515-521.

Schofield, Robert E. *The Enlightened Joseph Priestley: A Study of His Life and Work from 1773 to 1804*. University Park: Pennsylvania State University Press, 2004.

Schopenhauer, Arthur. *On the Freedom of the Will*. Trondheim: Royal Norwegian Society of Sciences, 1839.

Skorupski, John. *Why Read Mill Today?* London & New York: Routledge, 2006.

Smart, John Jamieson Carswell, and Bernard Williams. *An Outline of a System of Utilitarian Ethics*. Cambridge: Cambridge University Press, 1973.

Smart, Roderich Ninian. "Negative Utilitarianism." *Mind* 67, no. 268 (1958): 542-543.

Spir, African. *Denken und Wirklichkeit. Versuch einer Erneuerung der kritischen Philosophie*. Leipzig: J. G. Findel, 1873.

Taylor, Richard. "Fatalism." *The Philosophical Review* 71, no. 1 (1962): 56-66.

Taylor, Richard. *Action and Purpose*. Englewood Cliffs, NJ: Prentice Hall, 1966.

Urmson, James O. "The Interpretation of the Moral Philosophy of J. S. Mill." *The Philosophical Quarterly* 3, no. 10 (1953): 33-39.

Wilde, Norman. "Kant's Relation to Utilitarianism." *The Philosophical Review* 3, no. 3 (1894): 289-304.

Williams, Bernard. *Moral Luck*. Cambridge: Cambridge University Press, 1981.

Williams, Bernard. "Postscript." In *Moral Luck*, edited by D. Statman, 251-258. Albany: State University of New York Press, 1993.

Wood, Allen. *Kantian Ethics*. Cambridge: Cambridge University Press, 2007.

D. Menčik · *Conatus* 5, no. 1 (2020): 71-83
DOI: http://dx.doi.org/10.12681/cjp.22283

Identity Theft: A Thought Experiment on the Fragility of Identity

David Menčik
University of Novi Sad, Serbia
E-mail address: davidmencik@gmail.com
ORCID ID: https://orcid.org/0000-0003-3207-4795

Abstract

This paper intends to discuss some aspects of what we conceive as personal identity: what it consists in, as well as its alleged fragility. First I will try to justify the methodology used in this paper, that is, the use of allegories in ontological debates, especialy in the form of thought experiments and science fiction movies. Then I will introduce an original thought experiment I call "Who am I actually?," one that was coined with the intent to shed light on several aspects of the issue under examination, that is, the fragility of personal identity. Then I will move on to Christopher Nolan's film The Prestige, as well as to Derek Parfit's 'divided minds' thought experiment, to further discuss the fragility of personal identity; next to identity theft, the prospect of duplication is also intriguing, especially with regard to the psychological impact this might have on both the prototype and the duplicate. I will conclude with the view that spatial and temporal proximity or coexistence, especially when paired with awareness on behalf of the duplicates, would expectedly result in the infringement of the psychological continuity of one's identity.

Key-words: *social identity; physical identity; psychological identity; philosophy and fiction; divided minds; The Prestige; spatial and temporal coexistence*

I. Thought experiments, movies, and the debate on personal identity: Justifying the connection

Ever since the classical era there has been "a long-standing antagonism between poetry and philosophy."[1] Plato's idea of the philosopher is that of a person who has somehow managed to exit the cave, has seen the truth, and returns with the obligation to reveal to others the truth he has become aware of. In Plato's view the philosopher's task is to present things the way they truly are and persuade others by means of sound argumentation. Contrary to philosophy, poetry is only entitled to the

[1] Plato, *The Republic*, trans. Tom Griffith (Cambridge: Cambridge University Press, 2000), 607b.

subjective interpretation of the truth. Poetry and painting are seen by Plato as two forms of monoperspective representation prone to presenting things not as they truly are, but as they might, or might not be.[2]

This 'Platonic' antagonism between philosophy and art in general is still a valid standpoint. However, arguments in favor of this alleged dichotomy between philosophy and literature are being challenged by contemporary thinkers, and several scholars reject the thesis that such an opposition exists indeed. The basic arguments of the counter-dichotomy (or, counter-discontinuity) thesis are that:

> i. At least some works of literature present genuine and maybe even novel philosophical arguments that are of the same quality as the best of any bona fide philosophy.[3]
> ii. We have no reason to consider literature and philosophy to be dramatically different: (as they) both confront us with nonactual situations.[4]

Based on the counter-discontinuity approach one could make a correlation between philosophy and fiction, or, and this is exactly where this paper will focus, between thought experiments and science-fiction movies. Both thought experiments and sci-fi movies are indeed imaginative ways of illustrating and comprehending a problem. Damir Smiljanić suggests that especially issues related to "possible worlds, problems of identity and ethical dilemmas"[5] could be presented equally well by means of thought experiments and/or sci-fi books and movies, because both combine philosophical argumentation and narration.[6] One could come up with an argument against the discontinuity thesis by providing correlating examples from thought experiments and sci-fi movies that suggest possible answers to issues related to the fragility of identity. The fact that movies about identity-related issues are often filmed in

---

[2] The reference is to Book II of *The Republic*: "When a storyteller gives us the wrong impression of the nature of gods and heroes. It's like an artist producing the pictures which don't look like things he was trying to draw." Plato, *The Republic*, 377c.

[3] Bence Nanay, "Philosophy versus Literature? Against the Discontinuity Thesis," *The Journal of Aesthetics and Art Criticism* 71, no. 4 (2013): 352.

[4] Ibid., 351.

[5] Damir Smiljanić, "Upotreba misaonih eksperimenata u filozofiji i filmu," *Kultura* 142 (2014): 39 [translated from Serbian by the author].

[6] The view that thought experiments can be seen as a composition of argumentation and narration is similar to the one reached by Sören Häggqvist: "Thought experiments are not identical to arguments, they have to be seen as intimately connected to certain arguments." See Sören Häggqvist, "A Model for Thought Experiments," *Canadian Journal of Philosophy* 39, no. 1 (2009): 57.

such a way as to force the spectators to participate as puzzle-solvers, makes these movies even more philosophical in nature.[7]

Furthermore, identity-related movies can become an inspiration for philosophers to create engaging thought experiments.[8] No less significantly the storyline of a sci-fi book or movie can be placed in a different context, where the correlation between the problem presented in the movie is interwinned with a similar problem presented in the thought experiment, or another movie. Last but not least, the connection between thought experiments and movies is also valid due to that thought experiments are usually designed in such a way, as to hightlight the most important part of the issue that is under examination, exactly as movies are structured around the most important scene, that is sometimes repeated during the movie.

II. Who am I actually? A thought experiment on the possibility of social, physical and psychological identity change

Let's discuss the following – imaginary, though logically permissible – case.[9] Michael and Rodney are monozygotic twins, but their lives are completely different: Michael is successful, affluent, and has a family of his own; currently he is the CEO of a high-profile company. On the other hand, life hasn't been that kind to Rodney: he lives alone and is currently unemployed. What is most important for this thought experiment, is that Rodney and Michael *look* identical: *Their only physical dissimilarity is a tiny birthmark: a faint scar on one of Rodney's left foot toes* (1).[10]

On their birthday night Michael and Rodney go out to celebrate; after several hours of consuming large quantities of alcohol, Michael and Rodney have to decide whether they should drive back home, or take a taxi; Michael insists that he was sober enough to drive, and convinces Rodney to get into the car. This reckless decision of theirs proved fatal: a few minutes later their car crashed with another. The driver in the other car was killed on the spot; Michael got severely injured and was out cold, while Rodney only suffered

---

[7] See Warren Buckland, ed., *Hollywood Puzzle Films* (New York & London: Routledge, 2014). Also, Thomas Elsaesser, "The Mind-Game Film," in *Puzzle Films: Complex Storytelling in Contemporary Cinema*, ed. Warren Buckland, 13-41 (Oxford: Wiley-Blackwell, 2009).

[8] For example, the thought experiment I will present later on has been partly inspired by Damir Smiljanić's lectures on the Philosophy of Mind, during which he discussed extensively John Woo's film *Face Off*, and partly by Robert Louis Stevenson's novel *The Strange Case of Dr Jekyll and Mr. Hyde*.

[9] Thought experiments differ from scientific ones; they allow us to test our assumptions in non-factual situations. My thought experiment consists in a *possible* – yet, not *real* – course of events.

[10] The key steps of the thought experiment are highlighted in italics, and are also numbered.

minor injuries. Michael's injuries looked fatal to Rodney; convinced that his brother was going to die anyway, Rodney decides to steal Michael's identity and live his brother's – much more lucrative – life. *Rodney replicates his birthmark on Michael's toe, an identification mark that only Rodney had prior to this moment* (2). *Rodney switches documents and wears his brother's wedding ring on his finger* (3).

As soon as the ambulance arrives Michael is taken to the hospital, and Rodney tells the police that he, Rodney,[11] was the driver. Afterwards, Rodney and Marie (Michael's wife to whom Rodney presenst himself as his brother, Michael) visit Michael in the hospital; Michael's condition is critical, so the last think one would think of is to check for the faint scar on his toe. Michael has Rodney's identity card on him, and no wedding ring on his hand. Thus, everyone is convinced that it is Rodney the person who lies in bed heavily injured. This is the moment when the first identity change occurs; it is a change of *social identity*.[12] *Rodney presents himself as Michael, and injured Michael is presented as Rodney* (4).

Michael is in a comatose state; Rodney is worried that his tiny little birthmark could compromise the identity theft; therefore, he undergoes plastic surgery to eliminate his scar. Now only one person has the scar, Michael. Scar: Rodney's body; No scar: Michael's body. *When the scar is removed, Rodney's body looks as if it is Michael's, and Michael's as if it is Rodney's. This establishes a change in pseudo-physical identity* (5).[13]

After a couple of months Marie has already noticed that her husband's behavior (who in fact is Rodney) is not like it used to be. She suspects that the person who presents himself as Michael is actually Rodney. Marie decides to visit the comatose twin in the hospital, so as to check whether the person who is supposed to be Rodney has the tiny birthmark that makes the two brothers discernible. After she sees the scar she rests assured that the comatose patient is indeed Rodney; her line of reasoning is outlined as in premise (5): *Scar: Rodney; No scar: Michael*.[14]

---

[11] A man with Rodney's documents, that is, Michael.

[12] The term *social identity* is not easy to define. The closest to a definition could be this: social identity is *who one is* in relation to others, or who one *is presumed to be*.

[13] I use the greek word word *pseudo* as a prefix, so as to highlight the difference between *pretending* to be someone, *resembling* someone (physically), and *being* someone. For example, Rodney pseudo-physically *is* Michael, because he has changed his appearance so as to *look like* Michael. If, for example, the twins could exchange bodies or consciousness after the accident in some magical way, we would be justified to talk about (real) physical change and its impications.

[14] Marie's logic is flawed. According to the laws of deductive reasoning what she would only be justified to conclude would be that: Scar: Rodney; No scar: Not Rodney (modus tolendo tolens).

This instance is of key significance for the thought experiment, since it is at this point that the connection between social and physical identity is established. The very moment Marie sees the scar, she becomes absolutely convinced that the person laying in the hospital's bed is Rodney, and that the person who lives with her as her husband is indeed Michael (6).

While the real Michael is in a coma, his brother does his best to destroy everything that Michael has created. One day, though, a miracle happens: Michael unexpectedly wakes up. Rodney is the first to be informed; he rushes to the hospital, tells Michael that he (Rodney) had stolen Michael's identity, and proposes a trade off; since his trial for manslaughter after the accident was still in progress, and the jury would probably sentence in several years of imprisonment the drunk driver, that is, Michael (now he has woken up), Michael and Rodney both agree to switch identities once more. Rodney will take Michael's place and, if needed, do time instead of his brother; Michael, in turn, will compensate Rodney for an amount of one million dollars that will be transferred to Rodney's account. Michael will be free to go back to his wife and do his best to save his company.

It is important to note that this change of identity is only a social one; Rodney never tells Michael that Rodney has no scar, while Michael has. From this moment on Michael's social identity is 'Michael,' even though his pseudo-physical identity is 'Rodney,' and vice versa in the case of Rodney (7).

When Rodney goes to the bank to open the bank account where Michael would transfer the agreed amount, he decides to use his fingerprint as an identification; even identical twins have different fingerprints.

Rodney provides a physical proof for his identity, that confirms his social identity as Rodney. This adds a new dimension to the thought experiment. A physical identification mark that cannot be duplicated would prevent Michael from tricking Rodney and stealing his money (8).

This story could have a happy ending if Rodney hadn't been killed by his inmates shortly after he was found guilty by the jury and incarcerated. The tables were turned when Marie noticed the scar on Michael's left foot. It was the same scar she had seen on the hospitalized brother's toe in the hospital, that lead her to the conclusion that the person lying in bed was Rodney as shown in (6), though in fact it was Michael. Michael tells Marie that Rodney bears exactly the same scar; but when Rodney's corpse is being checked, Michael is shocked to see that there is no such scar on Rodney's leg. Marie is now pretty much sure that the person she lives with and presents himself as Michael is in fact Rodney who has stolen Michael's identity, and that the one who died in prison is actually Michael.

Rodney dies in prison; Marie notices the scar on Michael's leg; Rodney's

corpse has no scar. Marie concludes that Rodney, the twin with the scar, stole Michael's identity; Michael, the twin without the scar, has died in prison (9).

Because of this totally unanticipated course of events, Michael loses everything. When he goes to the bank, he is denied acess to the account due to the mismatch between his and Rodney's fingerprint. For all intents and purposes Michael is now socially Rodney because everybody *thinks he is* Rodney. At the same time, as far as the bank is concerned, he can't be Rodney, since his fingerprint doesn't match. Pseudo-physically Michael should be Rodney if the scar was to be the only determinant, but physically he is still Michael. Michael doesn't know who he is anymore; everything on him tell others that he is Rodney, but he knows he is not. In a moment of despair Michael loses it and kills himself. His last words left in a note were: "Who am I actually?"

Michael's fingerprint is different; socially he is both Rodney and not Rodney; pseudo-physically he is Rodney; Michael doesn't know who he is and kills himself (10).

III. Implications of the "Who am I actually" thought experiment and correlations with Christopher Nolan's *The Prestige* and Derek Parfit's 'Divided Minds'

a. Determinism and social identity changes

This thought experiment includes four changes in social identity. An issue that needs to be addressed is whether these changes occurred due to the free will of the persons involved, or were determined by factors external to the agents. It is obvious that the first change in both brothers' social identity is voluntary for Rodney, and involuntarily for Michael (3, 4). But still, even in the case of Michael, it cannot be deemed pre-determined. The identity exchange in (7), when the brothers decided to switch back again, has been voluntarily for both. This identity change is partially pre-determined and partially not, since it was step (3) that led to it, but the brothers didn't need to change their social identities back. The third major social identity change happens in (9), when Michael is socially restored as Rodney. This change is completely involuntary since it is determined by steps (1, 2, 3, 5, 6, 7), that is, all the steps that precede it, and include previous social identity changes, as well as pseudo-physical identity changes. Another important factor as far as the pre-determinated character of this identity change is concerned, is the fact that Marie reaches the wrong conclusion regarding Michael's physical identity while Michael was in a coma. The final social identity (10) is also involuntary and directly determined by (8), and indirectly determined by all previous steps. The fact that, as it is shown in the thought experiment, it is not

always us who decide whether our social identity will stay the same or not, leads to the question on why is social identity often as fragile as it has shown to be in the case of Michael.

b. The fragility of social identity, Michael's case, and the 'canary switch'

The thought experiment I presented above was intented to show that social identity can easily be altered. It may even seem that social identity is just an illusion. It is easy to fool other people in believing that one is the person one claims to be.[15] In Michael's case the fragility of social identity is shown in step (3), when his identity is stolen by Rodney by simple actions like exchanging documents and stealing a wedding ring. Furthermore, Michael's case shows that one can easily be mistaken for the person one is not (6, 9), and also that it may be not as easy to provide a solid proof of one's social identity in the case an identity theft occures (10).

The 'canary switch' scene in Christopher Nolan's film *The Prestige* presents masterfully the fragility of social identity. In the scene a canary is put in a cage on a table, and the cage is covered with a magician's scarf. The magician smashes the cage with his hand; then he removes the scarf and the cage as well as the canary have just vanished into thin air. Right after, the magician conjures the canary up safe in the scarf. The trick is performed twice during the film, and both times Nolan focuses on the impact it has on the spectators – it is awe and horror.[16] When the full version of the trick is exhibited, the truth is revealed: the table has a secret trap hatch, and right below the hatch is attached a compartment for the cage. The canary is being smashed indeed, and the trick lies in that it is artfully replaced by a 'twin' canary. What connects my thought experiment to Nolan's canary switch is the illusion of empirical singularity.

> An illusion involving the use of doubles functions through the dissimulation of the plurality of an object whereby the pledged

[15] One's social identity is not one's *possession*, nor an intrinsic attribute of one's essence. It is, after all, the only kind of identity that can be stolen, as it is shown in my thought experiment. For example, Rodney was able to steal Michael's social identity, just because he needed to convince other people that *he* is Michael. Fooling others was easy, since Michael and Rodney looked almost the same. On the other hand, Rodney could have never stolen Michael's *psychological identity*, because this kind of identity is inherent in one's essence as a part of one's consciusness.

[16] The feeling of *horror* emerges when one realizes what the trick consists in, as it happens when Nolan shows a boy having realized that the canary presented at the end of the trick is not the initial one, and asking: "But where's his brother?" Christopher Nolan, and Jonathan Nolan, *The Prestige: Screenplay* (London: Faber & Faber, 2006), 19.

> object does disappear or even die, but is immediately replaced by something else which looks exactly the same as the thing which was made to disappear.[17]

Though my scenario doesn't involve any disappearance trick, it nevertheless addresses the issue of the logical association between uniqueness (or, singularity) and duality, as the identity of both Rodney and Michael becomes dual at several instances. Marie, for example, believes that the birthmark on Rodney's toe is unique, and that it is therefore a proof that the person that bears it is Rodney. Rodney's identity can also be understood as dual because he is both Rodney (physically and psychologically) and Michael (pseudo-physically and ocassionally socially) at the same time. Rodney's duality consists in the fact that he is both himself as well as an impersonation of Michael, and the same applies to Michael, who is also an impersonation of Rodney (pseudo-physically and socially).

In my view all four types of personal identity I outlined in my thought experiment are interrelated.[18] The 'canary switch' scene in Christopher Nolan's film seems also capable of providing support to arguments in favor of the fragility of social identity. However, the fragility-of-identity hypothesis seems to be in need of further support; such support, I believe, could be sought in Derek Parfit's 'divided minds' thought experiment.

c. Psychological and physical identity: The originality of a replicated being

The problem of creating an identical copy of something that is presumed to be unique has fueled a persistent ontological debate that starts with Plato's discussion of *mimesis* and is present all the way through up to the critique of the 'cultrual industry' by Horkheimer. In "The Work of Art in the Age of Mechanical Reproduction" Walter Benjamin describes the mechanism of reproduction as follows:

> The technique of reproduction detaches the reproduced object from the domain of tradition. By making many reproductions it substitutes a plurality of copies for a unique existence.[19]

---

[17] Kwasu David Tembo, "On the Work of the Double in Christopher Nolan's The Prestige," in *The Cinema of Christopher Nolan: Imagining the Impossible*, ed. Jacqueline Furby, and Stuart Joy, 201-218 (London & New York: Columbia University Press, 2015), 203.

[18] Pseudo-physical identity and social identity are similarly structured, and are equally external to one's essence. I will argue that we could speak of four identities, provided that we keep in mind that the pseudo-physical identity doesn't exist per se; it requires a connection between social and physical identity. This identity emerges during our interaction with others (social identity) on the basis of how we look like (physical identity). This is why we may change our identity to look like someone else (pseudo-physical identity), so that other people will think that we are another person.

[19] Walter Benjamin, "The Work of Art in the Age of Mechanical Reproduction," in *Illuminations,*

In their *Dialectic of Enlightenment* Theodor Adorno and Max Horkheimer argue that the cultural industry has set out to reproduce everything that exists in the world: real-world objects are replicated by literature and the movie industry, and this results in the alienation from one's self, others, and real world problems. The distinction between the world of reality and the world of fiction is lost; people's attitude towards the world can be altered by regulating what they are furnished during their time of leisure, therefore controlling the media means enforcing conformity and manipulation.[20] In their works Adorno, Horkheimer, and Benjamin have discussed the devaluation of actuality in a world that relies on omnipresent reproduction. My focus here is not on political theory, of course; the issue that concerns originality and the way it is related to value, however, is definitely a common thread.

The question is whether creating identical copies of existing persons would compromise those persons' uniqueness or not. Leaving aside the various possible bioethical dilemmas that arise from human cloning, and limiting the focus on the issue of identity, there are two questions that need to be addressed:

a. Could my replica be physically identical to me?
b. Would the existence of a living replica of mine compromise my psychological identity?

The answer to the first question doesn't seem to be that complicated. If we had the power to create a machine that would produce replicas, like the one Tesla created for Angier in Nolan's *The Prestige*, I could be physically identical to my copy if 100% of my matter was used to give existence to my replica without destroying me at the same time.[21] However, it is important to note that even mere physical similarity, or should we say pseudo-physical identity between replicas and originals, could infringe psychological identity. We can see a reaction towards this infringement in Angier's expression after his first replica was created: Angier is stunned and horrified at the same time. Angier seems to be afraid that his own identity is being compromised by the

Essays and Reflections, ed. Hannah Arendt, 217-253 (New York: Schocken Books, 1999), 221.

[20] See Theodor W. Adorno, and Max Horkheimer, *Dialectic of Enlightenment*, trans. Edmund Jephcott (Stanford: Stanford University Press, 2002), 111-120.

[21] This, however, is neither *logically possible*, nor *technically feasible*. On the logical impossibility of creating two *indiscernible* beings see Leibniz's 'identity of the indiscernibles' priniciple, in Gottfried W. F. von Leibniz, *Philosophical Essays*, trans. Roger Ariew, and Daniel Garber (Indianapolis: Hackett, 1989), 41-42; on technical feasibility see Evangelos D. Protopapadakis, "Clones, Prototypes, and the Right to Uniqueness," *Agrafa – Journal of Philosophy of Psychoanalysis* 1, no. 2 (2013): 40-47, especially 42-44.

mere existence of a replica of his, even though Angier knows that it is only about a duplicate, and not the 'real him;' this is why he eventually decides to annihilate it. It seems that it is almost impossible to come to terms with the fact that I am staring at something that *is* and *is not* me at the same time.

Parfit's discussion is much more complex; in his 'divided minds' thought experiment Parfit discusses an imaginary situation in the context of which he is being 'teletransported' on planet Mars. During the process he loses consciousness and his body on Earth is being destroyed as scheduled, while at the same time a replicator on Mars produces a body that is absolutely identical with the one on Earth; his consciousness is also replicated: this new self has exactly the same knowledge, experiences and memories as the prototype.[22] Both identical beings (this on Earth, and that on Mars) share a common physical and psychological identity; the distance between them, however, made possible due to the duplicate's 'teletrasportation' and the destruction of the prototype, allow for the psychological impression that the replica is indeed a unique self. In this part the clone in Parfit's thought experiment doesn't share Angier's frustration due to his replication.

In the second part of the thought experiment Parfit enters the machine, but doesn't get teletransportated; he is being told, instead, that "the new scanner records your blueprint without destroying your brain and body."[23] An interesting question is whether he would consent to this if he was in advance informed that both versions of his self would coexist, albeit in different places. It is inconceivable for one to be in two places at the same time. "Wait a minute. If I am here, I can't also be on Mars."[24]

Shortly after, the prototype on Earth is told that his cardiac system has been damaged due to the replication process, and that he is going to die, while the replica on Mars will keep on living. When the original and the replica communicate, they are both convinced by their conversation that, even though physically identical, psychologically they are not the same person. "Call this the Branch-Line Case [...] though he is exactly like me, he is one person, and I am another."[25]

IV. A possible conclusion on the fragility of identity

In the first part of my article I set out to provide reasons for the methodology I used – both as far as my research, as well as the presentation of my results

[22] See Derek Parfit, *Reasons and Persons* (Oxford: Claredon Press, 1984), 199-203.

[23] Ibid., 198.

[24] Ibid.

[25] Ibid., 201.

are concerned. Even though the paper is quite short, and the focus has necessarily been narrowed down to only a few relevant cases, I believe that the conclusions I have reached and I am about to present could be of broader applicability and validity, eventhough just as hypothetical and speculative. Reaching speculative conclusions goes hand in hand with thought experiments, since they are just allegories intended not to prove, but to showcase. On the other hand, as far as identity isues are concerned, thought experiments is the best thing we can afford, since it is unethical to examine the fragility of personal identity by performing experiments on human beings.

Based on the discussion in this paper, one could reach two distinct conclusions. The first is that our social identity is not intrinsic to our essence; it almost sounds improper to speak of an identity at all, since social identity is entirely dependent on the objectification of one's own self by others. Perhaps it would be much more appropriate to understand one's social identity as *a means of interaction*. If persons didn't interact, there would be no such identity attached to one's self. We could imagine a situation in which a person, let's say, who sufferes from total amnesia is transferred in the wilderness. Living in the absence of any other person, this person is socially 'no-one,' as nobody else interacts with him. Next to this, as it has already been shown previously in this paper, one more argument in favor of the non-actuality of social identity is that it can easily be abolished, or transferred to another person.

The second conclusion would be that one's physical identity is interrelated with one's social and psychological identity. For example, similarity in physical appearance, or what I call *pseudo-physical identity*, may result in the social identification of two persons, as it is shown in the case of Rodney with regard to his scar, and in the 'canary switch' scene in Cristopher Nolan's *The Prestige*. Furhermore, as Parfit's thought experiment suggests clear no less than Angier's reaction to his replica, the psychological continuity of one's identity would be infringed if a. the prototype and the replica coexisted in spatial and temporal proximity, b. either of the two was aware of the existance of the other.

Acknowledgments

As the twin who survived, I didn't have the opportunity to experience how it is like to be a twin. It is probably due to this that issues as such attract my interest that much. But turning inclination into philosophical articulation is something that always requires the aid of mentors and collaborators. Therefore, I am indebted to Professor Damir Smiljanić for providing the initial and absolutely necessary inspiration and motivation for this paper. I am

also in Dr Goran Rujević's debt for being a critical reader and an extremely insightful commentator. I also need to thank Milena Stefanović who helped me discover my special interest in science fiction and encouraged me to make use of it in my philosophical inquiries. I am in Ruel Mannette's debt for providing insightful remarks and suggestions. Finally, I need to express my gratitude to both anonymous reviewers to whom my paper was assigned by the Editors of *Conatus* for their suggestions.

References

Adorno, Theodor W., and Max Horkheimer. *Dialectic of Enlightenment*. Translated by Edmund Jephcott. Stanford: Stanford University Press, 2002.

Benjamin, Walter. "The Work of Art in the Age of Mechanical Reproduction." In *Illuminations, Essays and Reflections*, edited by Hannah Arendt, 217-253. New York: Schocken Books, 1999.

Buckland, Warren, ed. *Hollywood Puzzle Films*. New York & London: Routledge, 2014.

Elsaesser, Thomas. "The Mind-Game Film." In *Puzzle Films: Complex Storytelling in Contemporary Cinema*, edited by Warren Buckland, 13-41. Oxford: Wiley-Blackwell, 2009.

Häggqvist, Sören. "A Model for Thought Experiments." *Canadian Journal of Philosophy* 39, no. 1 (2009): 55-76.

Leibniz, Gottfried W. F. von. *Philosophical Essays*. Translated by Roger Ariew, and Daniel Garber. Indianapolis: Hackett, 1989.

Nanay, Bence. "Philosophy versus Literature? Against the Discontinuity Thesis." *The Journal of Aesthetics and Art Criticism* 71, no. 4 (2013): 349-360.

Nolan, Christopher, and Jonathan Nolan. *The Prestige: Screenplay*. London: Faber & Faber, 2006.

Parfit, Derek. *Reasons and Persons*. Oxford: Claredon Press, 1984.

Plato. *The Republic*. Translated by Tom Griffith. Cambridge: Cambridge University Press, 2000.

Protopapadakis, Evangelos D. "Clones, Prototypes, and the Right to Uniqueness." *Agrafa – Journal of Philosophy of Psychoanalysis* 1, no. 2 (2013): 40-47.

Smiljanić, Damir. "Upotreba misaonih eksperimenata u filozofiji i filmu." *Kultura* 142 (2014): 34-51.

Tembo, Kwasu David. “On the Work of the Double in Christopher Nolan’s The Prestige.” In *The Cinema of Christopher Nolan: Imagining the Impossible*, edited by Jacqueline Furby, and Stuart Joy, 201-218. London & New York: Columbia University Press, 2015.

F. Peonidis · *Conatus* 5, no. 1 (2020): 85-100
DOI: http://dx.doi.org/10.12681/cjp.23604

Making Sense of Dignity: A Starting Point

Filimon Peonidis
Aristotle University of Thessaloniki, Greece
E-mail address: peonidis@edlit.auth.gr
ORCID ID: http://orcid.org/0000-0002-0109-4729

Abstract

Although appeals to human dignity became quite popular after the end of War World II in various moral and legal settings, the term retained an air of semantic indeterminacy, and scholars are of opposing minds concerning its usefulness and significance. In this essay I intend to offer a sketch of a "deflationary" account of human dignity – viewed as one moral value among many others – according to which it is conceived as the minimal respect we prima facie owe to our own personality, as well as to the personality of everyone else without any restriction or exception. This account is accompanied by a justification, which does not presuppose the endorsement of a particular moral theory, and envisages dignity as a bulwark to counter the morally abhorrent consequences of many categorical and normatively tainted dichotomies western societies have created.

Key-words: *dignity; self-respect; respect for personality; Immanuel Kant; offense; instrumentalization; discrimination; group inequality*

1. The understanding of moral values and principles is burdened by a long history, the different meanings they had in different historical periods, and the fact that most of them can be aptly characterized as "essentially contested concepts." As if these were not enough, any attempt to make sense of dignity as a moral value and give a plausible account of it has to deal with the diametrically opposing judgments concerning its role and significance that abound in recent literature.[1] On the one hand, there are

[1] Philosophical discussions start with the treatment of *dignitas* in Cicero's *On Duties*. It should be noted that the study of past conceptions of dignity is not only of historical interest, since they often re-surface in one way or another in relevant contemporary debates. For historical studies see Panajotis Kondylis and Viktor Pöschl, "Würde," in *Geschichtliche Grundbegriffe, Historisches Lexikon zur politisch-sozialen Sprache in Deutschland*, eds. Otto Brunner, Werner Conze, and Reinhart Koselleck, 637-677 (Stuttgart: Klett-Cotta, 1992); Marcus Düvell, Jens Braarvig, Roger Brownsword, and Dietmar Mieth, eds., *The Cambridge Handbook of Human Dignity: Interdisciplinary Perspectives* (Cambridge: Cambridge University Press, 2014); Martha C. Nussbaum, *The Cosmopolitan Tradition: A Noble but Flawed Ideal* (Cambridge, MA: Belknap

those who outright dismiss it as a useless normative concept, either because of its inherent ambiguity and indeterminacy, or because of its redundancy. In their view, there are other values that could more appropriately highlight and accurately describe a cluster of moral transgressions that are mistakenly dubbed as violations of dignity.[2] On the other hand, there are eminent philosophers, such as Jürgen Habermas, who confer dignity a highly elevated status and envisage it as a kind of supreme value, which

> performs the function of a seismograph that registers [...] those rights that the citizens of a political community must grant themselves if they are to be able to *respect* one another as members of a voluntary association of free and equal persons.[3]

In this essay I intend to explore the possibility of a middle-of-the-road, "deflationary" approach, which professes neither the abandonment of dignity nor the de-evaluation of a plurality of other crucial values *for the sake of* dignity. I am also interested in an approach that takes into account our pre-theoretical relevant intuitions, conceives dignity as a distinct value – one among many – that does not trespass on the scope of other values, and endorses a justification of it that does not rely on the wholesale acceptance of a particular comprehensive moral theory. I realize that this is a book-length project, but I would like to present here some preliminary basic remarks, that could serve as a starting point for further reflection.

2. On condition that we are seeking a conception of dignity that preserves its role as a distinct and clear value without overshadowing other equally basic values, we could start with the following description: *Dignity can be conceived as the prima facie moral claim to minimally respect our own personality and the personality of anyone else without any restriction or exception*. Given that

---

Press, 2019), 64-96; see also Remy Debes, ed., *Dignity: A History* (Oxford: Oxford University Press, 2017). For a detailed discussion of dignity as an essentially contested concept see Philippe-André Rodriguez, "Human Dignity as an Essentially Contested Concept," *Cambridge Journal of International Affairs* 28, no. 4 (2015): 743-756.

[2] See, for instance, Kondylis, "Würde," section VIII; Ruth Macklin, "Dignity is a Useless Concept," *British Medical Journal* 327, no. 7429 (2003): 1419-1420; Matti Häyry, "Another Look at Dignity," *Cambridge Quarterly of Health Care Ethics* 13, no. 1 (2004): 7-14, and Andrea Sangiovanni, *Humanity without Dignity: Moral Equality, Respect, and Human Rights* (Cambridge, MA: Harvard University Press, 2017), 13-71. The latter author takes great pains to show that the foundation of moral equality and basic rights does not depend upon the notion of dignity.

[3] Jürgen Habermas, "The Concept of Human Dignity and the Realistic Utopia of Human Rights," *Metaphilosophy* 41, no. 4 (2010): 469. Donelly and Griffin also belong to those who share the view that dignity grounds all basic human rights. See Jack Donelly, *Human Rights in Theory and Practice* (Ithaca and London: Cornell University Press, 2013), 29, 130-132; also James Griffin, *On Human Rights* (Oxford: Oxford University Press, 2008), 200-201.

this kind of treatment is due from anyone to everyone, dignity treats people as equals in this respect.

3. It has been argued that human dignity is subject to *variations* or *degrees* in the sense that people have more or less dignity depending on their character, their actions and omissions, and their way of life in general.[4] There is some confusion in this contention. Dignity as a distinct moral value should be kept apart from the plausible view that the moral assessment of others is conditioned by responsible agency on their behalf. The minimal respect for the personality of third parties dignity commands should be invariable and independent of any achievement, excellence, vice or malfeasance. In this sense, Dr. Pap (the Greek doctor who saved women from cervical cancer) and the worst criminal are equals as far as their dignity is concerned. Of course, this may not prevent us from having much higher esteem for Dr. Pap and show it on every occasion. We should not conflate dignity with personal *desert* in the wide sense and the endlessly debated obligations and duties ensuing from it.[5]

4. Failure to show the required respect for others' personality constitutes a form of moral harm that it is distinguished from other moral harms related to onslaughts on people's life, bodily integrity, well-being, freedom, and autonomy, or violations of various forms of distributive or non-distributive justice. I could degrade someone's dignity without causing her any other form of moral harm. I could also morally harm a third person without affecting her dignity. Let me give two examples from personal experience.

> A. When I was doing my military service, a non-commissioned officer ordered a small group of us to run around a courtyard half-naked and holding our rifles over our head under the excruciating Greek summer sun. This was not some kind of punishment for a petty offence we had committed, nor part of an officially approved training course. He just wanted to spite us for his own unfathomable reasons. Fortunately, as soon as we had started feeling the pain, a passing officer realized what was going on and ordered him to put an end to this absurd hazing. I cannot say that I suffered from this abuse of power any other moral harm apart from a personal humiliation.

---

[4] Suzy Killmister, *Contours of Dignity* (Oxford: Oxford University Press, 2020), 15.

[5] See among many others Serena Olsaretti, ed., *Desert and Justice* (Oxford: Clarendon Press, 2003). In this paragraph I rely on Darwall's distinction between "appraisal respect" and "recognition respect." Stephen Darwall, *The Second-Person Standpoint: Morality, Respect and Accountability* (Cambridge, MA: Harvard University Press, 2006), 122-126.

> B. A reckless driver was driving backwards in a one-way street and fell upon my parked car causing serious damage to it. It goes without saying that my property rights were violated and that my wellbeing was slightly affected, but I cannot maintain that he did not display the minimal required respect to my personality.

However, it is common for assaults on dignity to be accompanied with all sorts of moral harms. One of the most appalling examples one can think of is the treatment of the inmates in Nazi concentration camps. It would not be far from the truth to say that these individuals suffered any moral harm imaginable including a total and constant degradation of their personality and humanity.

5. How can dignity as described here be morally justified? A heuristic that was first used by the Stoics, and since then remains popular, consists in selecting one or more commonly held human features, which are supposed to make all humans bearers of a special value called dignity. In the Christian theological context, the dominant feature is man's likeness to God's image.[6] In a secular context, it is common for this purpose to single out one or more natural and/or cultural traits such as our ability to make crucial choices,[7] rationality, our supposedly unparalleled communication and interaction skills, humanity's brilliant record in the arts and the sciences,[8] etc. However, this approach is not devoid of problems. If one choses a single common feature of ours and argues that this confers dignity upon humankind, one has first to establish how common this feature is in the real world, and then explain why it is more significant than others. Why is rationality more appropriate than communication skills? Could not one equally argue for just the opposite? If one opts for a mixture of common features, one will keep wondering whether one's list can be ever finalized.[9] Moreover, one has to determine whether this feature (or features) possesses the necessary *normative* force to justify dignity, since we cannot derive values from facts. This is not an easy task, since it cannot be denied that most of these traits can be used – and have been used – for manipulating, humiliating and degrading our fellowmen. For instance, lying requires good communication skills and being instrumentally

---

[6] Kondylis, "Würde," section II.

[7] Giovanni Pico Della Mirandola, *On the Dignity of Man*, trans. Charles Glenn Wallis (Cambridge - Indianapolis: Hackett, 1998), 4-5.

[8] George Kateb, *Human Dignity* (Cambridge, MA: Harvard University Press, 2011).

[9] Cf. the difficulties surrounding the complete description of a person in terms of sufficient and necessary conditions as they are exposed in Dennett's seminal article "Conditions of Personhood," in *The Identities of Persons*, ed. Amélie Oksenberg Rorty, 175-196 (Berkeley, CA: University of California Press, 1976).

rational is a *conditio sine qua non* for succeeding in using someone as inferior for your own purposes. Not to mention our scientific achievements which allow us to turn the whole planet into a lifeless wasteland.

More promising seems the heuristic which aims to ground dignity on our capacity for moral agency, which is here regarded as a value in itself. This approach has two advantages. The first is that it avoids the naturalistic fallacy. The second is that by putting emphasis on the "capacity for moral agency" rather than "moral agency" it does not exclude from the scope of dignity minors, who are in the process of developing their moral capacities, and those adults whose capacity for doing the right thing has been impaired due to a variety of natural or social reasons. We can roughly discern two versions of it.

a. The Kantian version

Dignity does not have in Kant's moral universe the pivotal role of other concepts such as the moral law, the good will or the categorical imperative. It is briefly discussed in the second chapter of the *Groundwork of the Metaphysics of Morals* in which it is made clear that if x has dignity it cannot have a price, and thus it cannot be exchanged with something else. Rational beings (including humans) have dignity because of their capacity to legislate according to the moral law in a kingdom of ends and commit themselves to perform the moral duties this law prescribes. This capacity for universal and autonomous moral legislation gives human beings an elevated status and renders them superior to other natural creatures. Thus, according to Kant, the defining mark of an undignified person is one's contempt for this sort of demanding moral agency characterizing autonomous and rational beings.[10]

What is most troublesome with this justification of dignity is that it is part of a complex and comprehensive philosophical moral theory and it cannot be detached from its main body. It commits us to endorse the conception of a kingdom of ends populated by purely rational and abstract beings, the notion of the moral law, the idea of an absolutely good will, a particular account of how theoretical reason can be practical and other basic elements of Kant's moral discourse. This in turn implies that the justification sought can be formulated only within an exclusively Kantian philosophical context that is burdened with the well-known problems that concern the more theoretical part of the *Groundwork*.[11] Perhaps, someone might offer a

[10] Immanuel Kant, *Groundwork of the Metaphysics of Morals*, trans. Mary G. Gregor (Cambridge: Cambridge University Press, 1996), 4:434-4:440. See also Paul Sourlas, "Human Dignity and the Constitution," *Jurisprudence* 7, no. 1 (2016): 30-46, and Thomas E. Hill, "Kantian Perspectives on the Rational Basis of Human Dignity," in *The Cambridge Handbook*, 215-221.

[11] Cf. my *Autonomy and Sympathy: A Post-Kantian Moral Image* (Lanham, MD: University Press

charitable interpretation of Kant's ethics that can more or less accommodate these problems, but still the justification of dignity cannot be given in non-Kantian terms. In brief, if one seeks to dissociate dignity from inscrutable and ongoing exegetical controversies, the Kantian version cannot be her choice.

b. The non-Kantian version

Here one can invoke as a ground of dignity a wider and less demanding and sophisticated conception of moral agency that is not associated with a particular comprehensive moral theory.[12] Along these lines, moral agency may simply mean our ability to contain the pursuit of personal interests for the sake of others. Presumably, if people are so much self-absorbed that are completely indifferent to the good of others, they cannot show the minimal respect required by dignity. Moral agency is a prerequisite for honoring dignity. However, if we are to show that dignity is *grounded* in this type of moral agency, we have to establish that the latter *necessarily* implies the recognition of dignity.

Is this a tenable position? I think not. From the fact that people might care for the good of third persons we cannot infer that they regard them as their equals in the special sense required by dignity. Let us think of charity as it was practiced before the advent of the welfare state. Certain parts of the upper class who were aiding those at the bottom of the social ladder often entertained feelings of profound contempt for them. They believed that the poor were inherently inferior, since they were thought incapable of self-control and high feelings, as well as of cultivating the excellences of the ruling class. The desire to act for one's benefit does not necessarily imply respect for one's dignity.[13] It is consistent with insulting, humiliating and manipulating the object of our moral concern. Thus, special moral argumentation should be adduced for the justification of dignity and this leads us back to square one.

Let me explore another route. In the history of western civilization, there are many examples of well-entrenched categorical normative dichotomies that divide people to superior and inferior based on their natural and/or social characteristics. The composition of the groups formed by each dichotomy is on certain occasions subject to change, since there might be upward and downward mobility, but this does not affect the moral, financial, legal, political, and social inequalities that exist between the two groups. From

---

of America, 2005), 28-32.

[12] Cf. Griffin, *Human Rights*, 200.

[13] At the same time, we can think of people who are so committed to the wellbeing of others and so eager to work for their benefit, that they end up compromising voluntarily or involuntarily their own dignity.

classical antiquity the moderns have inherited the distinction between men and women, freemen and slaves, rich and poor, nobles and commoners, civilized and barbarians, the powerful and the powerless; from Christianity the distinction between the faithful and the infidels as well as between the orthodox and the heretics; from European colonial expansionism the distinction between whites and non-whites; from the rise of the nation-state the distinction between compatriots and aliens; from the development of medicine and its institutionalization the distinction between the healthy and the sick and so forth.

However, gradually the propriety of these dichotomies started being challenged in theory and in practice. The dismal consequences they usually had for those who were cut off from the dominant groups could no longer be accepted or tolerated by certain groups or parts of the body politic. There was a moral reaction to the injustices perpetuated by these categorical divides, which, despite certain occasional failures or extremities, served as a justification for a variety of public interventions. This onerous and painful process has started taking shape with the two revolutions of the late 18th century and it continues until now with remarkable results. Slavery was abolished and the term "barbarian" is now used only by historians and in inverted commas. The establishment of representative democracy and universal franchise empowered common people at the expense of nobility, while religious freedom saved "infidels" and "heretics" alike from indescribable hardships. The poor enjoy now a better quality of life thanks to the welfare state and the introduction of certain redistributive schemes. Men and women are officially regarded as politically and legally equal, although we cannot claim that we have won the war against sexism. In a similar vein, racism has received many blows, but it has not been defeated yet. Supranational political institutions, such as the European Union, have done great progress in reconciling European nations, although we are far from forging a shared European identity. Finally, the reaction against the dominance of medical discourse and its constant intervention in people's lives gave rise to the postwar movement for the rights of patients, contributed to the improvement of conditions in mental institutions and helped changing our perceptions towards certain forms of sexual behavior, which in the past were wrongly classified as diseases.

The abolishment of these distinctions and the eradication of the harms they generate where possible, the reduction of the existing gaps and in all cases the re-evaluation of the supposed normative categorical differences between superior and inferior groups – or to put it bluntly realizing that whites are not better than non-whites, men than women, or the English than the Irish – continue to be an objective of mainstream contemporary moral and political theory, although it is doubtful whether in an era where so many

things are said and written by so many people it could be as influential as it was in the past. If there is moral progress, it depends heavily on the extent to which this objective can be achieved. However, there is no royal road to this destination as attitudes do not change easily, people tend to invent novel normatively tainted categorical dichotomies and there is no guarantee that our moral achievements will last forever.

Dignity as it has been described can be seen as a moral value, which is justified by its role in the struggle against the above categorical dichotomies and the harms suffered by those who are not favored by them. It urges us to react against these divisions – or not to create new ones – by showing the necessary self-respect and by minimally respecting the personality of our fellowmen irrespective of who they are or what they have done. In this manner, it *ascribes* people a basic form of moral equality concerning the treatment of their personality that functions as a bulwark against the dreadful hierarchical and discriminatory relations and the illegitimate exclusions arising from the above divides our civilization has abundantly produced. Undoubtedly, the pursuit of dignity does not suffice to bring about the desired result, given that a whole cluster of values and attitudes should come in to play, but its role is distinct and not at all negligible. In addition, this way of understanding the value of dignity does not commit us into endorsing a particular comprehensive philosophical moral theory.[14]

6. Shifting now to a more mundane level it is helpful to identify particular types of action that could count as violations of dignity as it is understood in this essay.[15] Providing a complete list would be unfeasible, but if we focus on acts that degrade the personality of others either by *offending* them or by using them as mere *instruments*, it would seem that we are on the right track. In the first category we can include rudeness,[16] hate speech, racist, sexist and xenophobic behavior, various humiliating practices, bullying, hazing, sexual harassment, forms of legal punishment that aim at degrading the offender,[17]

---

[14] This particular understanding of dignity and its justification is incompatible with the view that dignity is the founding principle of all rights. As Rosen correctly points out, the right to respect one's dignity is the only right our conception of dignity could justify. See Michael Rosen, *Dignity: Its History and Meaning* (Cambridge, MA: Harvard University Press, 2012), 62.

[15] See Killmister, *Contours*, chapters 3 and 4, and J. M. Bernstein, *Torture and Dignity: An Essay on Moral Inquiry* (Chicago and London: The University of Chicago Press, 2015); in both works there are in depth and nuanced discussions that shed light on how assaults on dignity can be carried out.

[16] A rude person is someone who does not show respect for the personality of others par excellence.

[17] These are rare now, but they have not been eliminated. To give an example from Greece, statutory decree no 4000/1958, which was in force until 1983, allowed the public humiliation of youngsters who were found guilty of certain forms of provocative and offensive behavior.

vulgar satire, the disrespectful treatment of hospital patients, and so forth. The second category contains types of action where agents use other people as instruments according to their desires and interests.[18] In all these cases, they assume a position of superiority, which makes them oblivious of their duty to regard others as equals regarding the minimal respect for personality owed to them. Here we can think of lying, deception and manipulation,[19] torture, rape, practices of exploitation, modern forms of slavery like human trafficking etc. As said before in many of these cases assaults on dignity coexist with causing other moral harms of various sorts. However, a caveat is in order here. The moral depravity of these affronts on dignity does not render them *ipso facto* legally punishable. Even if we agree that prostitution violates the dignity of sex workers, this cannot be the only reason for prohibiting it.

7. As said before, dignity as minimal respect for personality is something that is *prima facie* owed to everyone without any prior restriction or exception. However, at which age human beings can be regarded as bearers of dignity? I would say from the moment they come into this world and become subject to many of the aforementioned categorical dichotomies. We are not allowed to humiliate or degrade a newborn baby because of its sex or its color or to use it merely as an instrument for our own purposes. It is reasonable to assume that the harm resulting from violations of dignity is felt more easily and intensively by subjects whose personality has been developed, but this does not affect our relevant obligations. Moreover, in certain occasions the harm done may be felt with delay, since the subject might have had not understood the violation the moment it took place and have had taken cognizance of it much later.

Does it make any sense to say that we can show disrespect for the "dignity" of embryos or *in vitro* fertilized human eggs? Neither the history of the term, nor the justification offered here, nor its everyday use condone the ascription of dignity to unborn human beings. To speak of the personality of an embryo in the sense that personality is being discussed in this essay would be a category mistake. The harms we can impose on unborn human beings are open to discussion and are often acrimoniously debated but appeals to their dignity just add to confusion.

---

[18] The above description retains a high level of generality. A more detailed approach could produce a classification of violations of dignity of this type drawing on Margalit's scheme according to which instrumentalization might mean treating people (a) as objects, (b) as animals, (c) as machines or (d) as subhumans. Avishai Margalit, *The Decent Society*, trans. Naomi Goldblum (Cambridge, MA: Harvard University Press, 1996), 89.

[19] It should be noted at this point that some cases of manipulation are regarded morally justifiable all things considered. A standard example is A's decision to deceive B in order to save C's life. This does not mean that B's dignity has not been compromised. It is just that in A's best judgment the moral gains from saving C's life outweigh the losses resulting from B's violation of dignity. That's why our obligation to show self-respect, and also to respect the personality of third parties is a prima facie obligation.

8. Some adults may not be affected by assaults on their dignity either because they do not realize what is going on, or because they are used to them. For example, in the Stoic tradition agents are urged to acquire sufficient control on their emotions so as to counteract the adverse effects of offense. Gaius Musonius Rufus, a Stoic philosopher who lived in the 1st century AD, argued that allowing yourself to suffer by insults is a token of pettiness, while remaining undisturbed in the face of them is a sign of magnanimity.[20] However, the manner people react to violations of their dignity does not affect our obligation to honor it. From the fact that they may decide not to react for a variety of reasons to these violations it does not follow that they waive their expectations of being treated with respect.

9. From the description given it follows that individuals can compromise their own dignity. This idea is also incorporated in everyday morality as it is attested by the fact that we expect others to behave in a dignified matter throughout their whole life. This means among others that we want them to show self-respect, be sincere and straightforward, not act in ways that make others pity them when they can do otherwise, avoid being obsequious and be vigilant about not ridiculing or humiliating themselves. We also realize that this is not something easily accomplished especially "in times that try men's souls." Admittedly, there are differences in the way agents understand self-inflicted violations of dignity, but the following passage from Kant, although it was written a long time ago, expresses some of the relevant intuitions at least of those sharing a secular conception of morality.

> Be no man's lackey. Do not let others tread with impunity on your rights. Contract no debt for which you cannot give full security. Do not accept favors you could do without, and do not be a parasite or a flatterer or (what really differs from these only in degree) a beggar. [...] Kneeling down or prostrating oneself on the ground, even to show your veneration for heavenly objects is contrary to the dignity of humanity, as is invoking them in actual images; for you then humble yourself, not before an ideal presented to you by your own reason, but before an idol of your own making.[21]

How should we react when we see adults compromising their own dignity regardless of any other moral harm they might cause to themselves and third

[20] Gaius Musonius Rufus, *Lectures and Sayings*, trans. Cynthia King (Scotts Valley, CA: Create Space, 2011), 50-51.

[21] Immanuel Kant, *The Metaphysics of Morals*, trans. Mary G. Gregor (Cambridge: Cambridge University Press, 1996), 6: 436-437.

parties? We should take into consideration that they violate only a duty to themselves. This implies that we can warn them about the moral significance of their actions (of which they might be unaware) and give counsel, but we are not allowed to stop showing the minimal owed respect for their personality. These individuals have not breached some kind of contract or a binding agreement they made with us. Therefore, we are not free from our obligation to treat them in the same way we treat persons who show self-respect. Two wrongs do not make a right. Besides, we have to bear in mind that these persons might have found themselves in such adverse circumstances that they were compelled to sacrifice their dignity for the sake of other overarching values.

10. It is commonly held that there is a close connection between the protection of human life and the protection of dignity. For instance, one of the aims of UNESCO's *Universal Declaration on Bioethics and Human Rights* (2005) is "to promote respect for human dignity [...] by ensuring respect for the life of human beings," while Killmister points out that dignity "grounds the value of human life."[22]

Claims like these are plausible but they still should be qualified. They fail to take into account the possibility of dying to avoid an undignified life or the prospect of an undignified life. In extreme circumstances, people may decide that death is preferable to a life in which their dignity is (or will be) constantly violated and there is no way out of the horns of the dilemma they face. One could think of the scores of women who committed suicide to avoid becoming spoils of war and in the hands of victors or of many public officials who put an end to their life to spare the embarrassment and humiliation of a public trial and subsequent imprisonment. We could also think of seriously ill patients who do not want to live any more an undignified life with no prospect of recovery. Judging the decisions of these people draws heavily on the views about the morality of suicide, medically assisted suicide or euthanasia one holds, and this is not the place to discuss such controversial issues. My point is simply that respecting my dignity does not *necessarily* imply respecting my life, and that respecting my life does not *necessarily* imply respecting my dignity.

11. Surveying the history of the understanding of the value of dignity from Cicero to Kateb, one cannot but observe a common theme that is not missing from most accounts: dignity is something that man possesses and animals lack. This makes us superior to other natural creatures. Can we accept this claim as easily as our predecessors did? It cannot be disputed that our conception

[22] Suzy Killmister, "Dignity: Not Such a Useless Concept," *Journal of Medical Ethics* 36, no. 3 (2010): 164.

of the normative value of the animal world has been substantially improved. Bentham introduced the idea that we have an obligation not to inflict pain on them and Kant first pointed out that being cruel to animals is a sign of wickedness, but, only recently, certain forms of maltreatment of animals became in many countries an offense punishable by law. Some philosophers argue that these developments do not suffice, and that we should impose a general banning on the killing of animals of any kind. In this changing context, the question about the dignity of animals should be posited again. One could turn to the biological differences between humans and animals, which reveal a completely different picture from the one most philosophers had in the past, but any effort to formulate a novel conception of dignity that is not confined to *homo sapiens* will run into considerable difficulties. Our efforts to seek one or more indisputably common characteristics to ground such a conception of dignity are likely to prove at least as troublesome as our effort to ground human dignity in our characteristics.[23] Is it possible to construct a conception of animal dignity that would be *different* from human dignity? In a deplorable zoo that the Municipality of Thessaloniki still refuses to shut down, a wolf was being kept in captivity in a small enclosed space. The poor animal kept endlessly moving in circles, which resulted in a ditch as deep as to hide the largest part of the wolf's body. It was a disheartening sight. It occurred to me that, irrespective of any other obligations we might have to animals, we show proper respect for their dignity if (a) we do not detach them from their natural environment and (b) we do not make them behave in ways not befitting their nature. This idea in my view could give an edge in the efforts to promote animal welfare, and it should be further explored.

12. Up to now, I have been arguing on the not unanimously held assumption that dignity is a *moral* value.[24] The question now is how the moral conception I put forward or some other conception affects the law. A cautionary approach

[23] Perhaps, one could try to ground this common conception of dignity on De Waal's theory according to which primates have certain primitive moral emotions, such as sympathy, which are encountered in more evolved and complex forms in humans. However, it can be objected that the differences between their and our moral emotions are so big (not to mention non-primates) that they render this endeavor unworthy of the effort. See Frans de Waal, *Primates and Philosophers: How Morality Evolved* (Princeton, NJ: Princeton University Press, 2006).

[24] Waldron, for instance, claimed in a well-known work that the notion of dignity makes sense only as a legal value. However, the arguments he uses to show the impossibility of a purely moral approach, such as the problems surrounding the Kantian conception of dignity or the existence of racist moral philosophers in the past, are not strong enough to justify his conclusion. Moreover, he tends to see in many legal rules and provisions a strong concern for the protection of dignity, which is not obvious to me. For instance, the maxim that all defendants are entitled to legal representation is better justified by law's concern to protect the freedom and the welfare of its subjects and to avoid injustice rather than to honor their dignity. See Jeremy Waldron, *Dignity, Rank and Rights* (Oxford: Oxford University Press, 2012), 13-78.

is here appropriate. Presumably, legislators can incorporate references to dignity in official legal texts, a practice that is common after the end of War World II. What they cannot incorporate are long philosophical essays explaining in detail the manner they understand it. Consequently, dignity in legal texts retains its ambiguity and as it has been aptly remarked this semantic indeterminacy can pave the way to novel and more comprehensive and sophisticated philosophical explorations of the term, but "when dignity is elevated from its status as a moral value to that of a judicially enforceable legal rule, its ambiguity is less a virtue than a liability."[25] It is not accidental that the Supreme Court of Canada in 2008 abandoned the use of dignity as a legal test in cases of violations of equality rights as inappropriate for the administration of justice.[26] Does this mean that dignity should have no place in legal systems?

Not necessarily. Ordinary legislators when they decide which forms of human behavior should be within the scope of criminal law, they could take into account actions that violate human dignity (understood in a strictly defined sense) and inquire whether these violations are so serious as to be accompanied by legal sanctions. When, for instance, the form of punishment for rape is determined, legislators will have to take into consideration that rape constitutes among other things a serious assault on woman's dignity.

[25] Thomas M. J. Bateman, "Human Dignity's False Start in the Supreme Court of Canada: Equality Rights and the Canadian Charter of Rights and Freedoms," *The International Journal of Human Rights* 16, no. 4 (2012): 577. Let me give an example of a controversial judicial decision involving the protection of dignity. In 2006, the German Federal Constitutional Court reached a landmark decision. It stroke down as unconstitutional §14.3 of the Aviation Security Act, which authorized the Federal Government to shoot down passenger planes that had been hijacked by terrorists with the intention to crash them down causing multiple casualties. The Court invoked among others the first article of the German Constitution (Basic Law) that safeguards the inalienability of human dignity. According to the Karlsruhe judges this provision makes it "absolutely unthinkable to legislate the purposeful taking of the life of persons who are in such a desperate position," that is, the passengers of the hijacked plane. Contrary to the Court's ruling, one could argue that in such cases the innocent passengers are doomed anyway, since they can do practically nothing to overcome the terrorists and regain control of the plane. Taking this into account, which of the following two possible decisions would better respect the dignity of those unfortunate passengers? According to the judges we should allow them to be used (along with the plane) as tools in the hands of terrorists to bring havoc. According to the opposite view, we respect better the passengers' dignity by sacrificing them in order to save the lives of many more unsuspected people and thwart the terrorists' plans. This also allows us to honor them posthumously as martyrs and heroes, something that cannot be done if we opt for non-interference. If I had to decide what I would like to happen to me – and discounting temporarily any harm I may inadvertently cause to third parties – I would prefer the second option because dignity as self-respect includes my current desire that others may have a justifiably positive opinion of me after my death. For a synopsis of the court's decision see https://www.bundesverfassungsgericht.de/SharedDocs/Pressemitteilungen/EN/2006/bvg06-011.html.

[26] R v Kapp 2008 SCC 41, https://scc-csc.lexum.com/scc-csc/scc-csc/en/item/5696/index.do.

Thus it would be easier for ordinary judges to apply rules of the type, say, "rape which is defined as any type of sexual penetration without the consent of the victim is punished as follows" rather than rules of the type "violations of dignity are punished as follows." I am not saying that legislators have an easy job to do or that their judgment would be infallible and immune to criticism, but in this manner the danger of using dignity in courts as an elastic concept which allows for a wide array of meanings is minimized.[27]

13. In this essay I set out to take down certain thoughts and ideas that could be used to defend a plausible conception of dignity viewed as a distinct moral value. In my view, the philosophical endeavors to make sense of dignity should avoid its underestimation or overestimation, not render its endorsement dependent on the wholesale acceptance of a particular moral theory and take into account some of our relevant pre-theoretical intuitions. I do not claim that my arguments suffice for constructing a full-blown theory of dignity, the final form of which is unknown to me, but I believe they could serve as raw materials for starting to work on it.

Acknowledgements

I would like to thank Nicholas Vroussalis and the two anonymous referees for valuable criticisms and suggestions.

References

Bateman, Thomas M. J. "Human Dignity's False Start in the Supreme Court of Canada: Equality Rights and the Canadian Charter of Rights and Freedoms." *The International Journal of Human Rights* 16, no. 4 (2012): 577-597.

Bernstein, J. M. *Torture and Dignity: An Essay on Moral Inquiry*. Chicago and London: The University of Chicago Press, 2015.

Darwall, Stephen. *The Second-Person Standpoint: Morality, Respect, and Accountability*. Cambridge, MA: Harvard University Press, 2006.

De Waal, Frans. *Primates and Philosophers: How Morality Evolved*. Princeton, NJ: Princeton University Press, 2006.

[27] Admittedly, this suggestion does not apply in cases in which the promotion or protection of dignity is prescribed by the constitution [Constitution of the Republic of South Africa (article 10), German Basic Law (article 1, §1)]. Here judges are free to invoke the relevant constitutional provisions, but the establishment of a common understanding of dignity by courts will reduce the risk of arbitrary judicial interpretations of this value.

Debes, Remy, ed. *Dignity: A History*. Oxford: Oxford University Press, 2017.

Dennett, Daniel. "Conditions of Personhood." In *The Identities of Persons*, edited by Amélie Oksenberg Rorty, 175-196. Berkeley, CA: University of California Press, 1976.

Donelly, Jack. *Human Rights in Theory and Practice*. Ithaca and London: Cornell University Press, 2013.

Düvell, Marcus, Jens Braarvig, Roger Brownsword, and Dietmar Mieth, eds. *The Cambridge Handbook of Human Dignity: Interdisciplinary Perspectives*. Cambridge: Cambridge University Press, 2014.

Griffin, James. *On Human Rights*. Oxford: Oxford University Press, 2008.

Habermas, Jürgen. "The Concept of Human Dignity and The Realistic Utopia of Human Rights." *Metaphilosophy* 41, no. 4 (2010): 464-480.

Häyry, Matti. "Another Look at Dignity." *Cambridge Quarterly of Health Care Ethics* 13, no. 1 (2004): 7-14.

Hill, Thomas E., Jr. "Kantian Perspectives on the Rational Basis of Human Dignity." In *The Cambridge Handbook of Human Dignity: Interdisciplinary Perspectives*, edited by Marcus Düvell, Jens Braarvig, Roger Brownsword, and Dietmar Mieth, 215-221. Cambridge: Cambridge University Press, 2014.

Kant, Immanuel. *Groundwork of the Metaphysics of Morals*. Translated by Mary J. Gregor. Cambridge: Cambridge University Press, 1996.

Kant, Immanuel. *The Metaphysics of Morals*. Translated by Mary J. Gregor. Cambridge: Cambridge University Press, 1996.

Kateb, George. *Human Dignity*. Cambridge, MA: Harvard University Press, 2011.

Killmister, Suzy. "Dignity: Not Such a Useless Concept." *Journal of Medical Ethics* 36, no. 3 (2010): 160-164.

Killmister, Suzy. *Contours of Dignity*. Oxford: Oxford University Press, 2020.

Kondylis, Panajotis, and Viktor Pöschl. "Würde." In *Geschichtliche Grundbegriffe, Historisches Lexikon zur politisch-sozialen Sprache in Deutschland*, edited by Otto Brunner, Werner Conze, and Reinhart Koselleck, 637-677. Stuttgart: Klett-Cotta, 1992.

Macklin, Ruth. "Dignity is a Useless Concept." *British Medical Journal* 327, no. 7429 (2003): 1419-1420.

Margalit, Avishai. *The Decent Society*. Translated by Naomi Goldblum. Cambridge, MA: Harvard University Press, 1996.

Musonius Rufus, Gaius. *Lectures and Sayings*. Translated with an introduction by Cynthia King. Scotts Valley, CA: Create Space, 2011.

Nussbaum, Martha C. *The Cosmopolitan Tradition: A Noble but Flawed Ideal*. Cambridge, MA: Belknap Press, 2019.

Peonidis, Filimon. *Autonomy and Sympathy: A Post-Kantian Moral Image*. Lanham, MD: University Press of America, 2005.

Pico Della Mirandola, Giovanni. *On the Dignity of Man*. Translated by Charles Glenn Wallis. Cambridge - Indianapolis: Hackett Publishing Co., 1998.

Olsaretti, Serena, ed. *Desert and Justice*. Oxford: Clarendon Press, 2003.

R. v. Kapp 2 S.C.R. 483, 2008 SCC 41.

Rodriguez, Philippe-André. "Human Dignity as an Essentially Contested Concept." *Cambridge Journal of International Affairs* 28, no. 4 (2015): 743-756.

Rosen, Michael. *Dignity: Its History and Meaning*. Cambridge, MA: Harvard University Press, 2012.

Sangiovanni, Andrea. *Humanity without Dignity: Moral Equality, Respect, and Human Rights*. Cambridge, MA: Harvard University Press, 2017.

Sourlas, Pavlos. "Human Dignity and the Constitution." *Jurisprudence* 7, no. 1 (2016): 30-46.

Waldron, Jeremy. *Dignity, Rank and Rights*. Oxford: Oxford University Press, 2012.

I. Wolfe & M. Guiahi · *Conatus* 5, no. 1 (2020): 101-108
DOI: http://dx.doi.org/10.12681/cjp.22528

Balancing Conscience: A Response to Fernandes & Ecret*

Ian Wolfe
Children's Mercy Bioethics Center, USA
E-mail address: wolfe370@umn.edu
ORCID ID: http://orcid.org/0000-0003-3308-3749

Maryam Guiahi
University of Colorado, USA
E-mail address: maryam.guiahi@ppcentralcoast.org
ORCID ID: http://orcid.org/0000-0002-6005-8314

Abstract

There are many lessons that bioethics can learn from the Holocaust. Forefront are the lessons from the Nuremberg trials and the formation of research ethics. An often-overlooked lesson is how the Nazi regime was able to construct a hierarchy in such a way that influenced people to act in horrendous ways. Fernandes & Ecret, writing in Conatus – Journal of Philosophy 4, no. 2 (2019), highlight the influence of hierarchy on the moral silence of nurses and physicians within the Nazi regime. While we greatly enjoyed the paper, and think it is an important contribution, we find several misrepresentations of current bioethical discourse. There is not a global acceptance of euthanasia or medical aid in dying, the contemporary position in bioethics does not favor removal of conscience-based protections, and the lack of personal conscience-based protections was not the main factor in the analysis of Nazi medical hierarchy. The authors' overall conclusion that their analysis suggests the importance of strengthening personal conscience-based objections to prevent medical hierarchy from influencing immoral behavior misses the more significant issue of institutional behavior. Instead, we argue, that the lessons from the analysis of Nazi nurses and physicians speak to the importance of protecting patients from institutional conscience-based objections that violate patient rights of access to legal medical services. This paper will respond to the misrepresentations. We highlight the growing threats to health care access from religious affiliated institutions, the threats to professional ethics and physician and nurse scope of practice. We conclude that the analysis by Fernandes & Ecret does point out an important lesson, but rather than showing the importance of individual conscience protections, it speaks to the importance of protecting patient rights.

Key-words: *conscientious objection; patient rights; bioethics; organizational ethics; professional ethics*

*A. Fernandes, and D. Ecret, "The Effect of Hierarchy on Moral Silence in Healthcare: What Can the Holocaust Teach Us?" *Conatus - Journal of Philosophy* 4, no. 2 (2019): 21-43.

We read with great interest "The Effect of Hierarchy on Moral Silence in Healthcare: What Can the Holocaust Teach Us?" by Fernandes & Ecret, which presents a powerful and important discussion of hierarchy and moral silence in healthcare.[1] In the article, the authors highlight the dangers of silence in the face of ethical violations, encourage the moral voices of nurses and physicians, and believe the lessons of the Shoah should be better materialized in professional education. We agree!

The authors primarily hypothesize how ingrained hierarchy had a role in shaping moral actions during the Holocaust. They are right that it is foolish to ignore these lessons, giving way to the complacency that it couldn't happen here or now. We agree that hierarchy of the Nazi regime worked to ensure physicians and nurses stayed silent and were complicit. We, however, challenge the authors' concluding assumptions that:

> i. there is a global acceptance of euthanasia and assisted suicide in the "medical profession" for conditions such as depression, schizophrenia, autism, addiction, and transgenderism.
> ii. "contemporary" literature in bioethics favors removal of conscience protection laws.
> iii. lack of individual conscience protections was the main factor in complicit behavior in the Shoah.

The authors conclude that these assumptions support the argument for strengthening individual provider conscience protections against the dangers of medical hierarchy, preventing moral silence. Ensuring protections are important. The more prescient threat is not medical hierarchy but religious affiliated organizations holding institutional conscience-based refusals against offering medical services.

Hospitals that hold restrictive conscience-objections prohibit staff from providing or discussing care with their patients. The US government, instead of strengthening patient protections, is seeking to expand protections for those entities.[2]

We will address our criticism of the three assumptions that Fernandes & Ecret base their conclusion on. We will then show how the lessons

---

[1] Ashley K. Fernandes, and DiAnn Ecret, "The Effect of Hierarchy on Moral Silence in Healthcare: What Can the Holocaust Teach Us ?" *Conatus - Journal of Philosophy* 4, no. 2 (2019): 21-43.

[2] Ian D. Wolfe, and Thaddeus M. Pope, "Hospital Mergers and Conscience-Based Objections – Growing Threats to Access and Quality of Care," *The New England Journal of Medicine* 382 (2020): 1388-1389.

from Fernandes & Ecret's analysis seem to point towards the dangers of an institution setting a specific conscience and overriding clinician conscience and ethics against the rights of patients.

Assumption I

Assisted suicide, now referred to as medical aid in dying (MAID), is fairly accepted in some medical professions, but there is by no means "global" acceptance.[3] Although MAID is certainly gaining momentum, it is legal in only seven US jurisdictions, for limited indications, and a rigorous process for access.[4] The American Medical Association still considers it "fundamentally incompatible with the physician's role."[5] The American Nurses Association only recently changed their policy around MAID that recognizes it as a legal and ethical right, highlights the obligation of the nurse to support their patient in that choice and through the process, but still maintain that nurses are ethically prohibited from administering medications for that purpose.[6] Further, euthanasia, which is different and distinct from MAID, remains illegal in the United States, and most other countries. One or two instances in history do not equate to accepted practice.

Assumption II

The authors' claim that contemporary bioethical literature supports removal of conscience protection laws is false and is a mischaracterization of this debate. The authors misrepresent Stahl & Emmanuel's argument by claiming it called for an end to conscience protection laws.[7] The argument Stahl & Emmanuel make is that a clinician should not be allowed to utilize conscience protection laws to violate a patient's right to access of a legal and medically accepted treatment, particularly when the clinician is federally funded or practices in an organization that receives public funds. Patient rights must be the primary concern. Clinicians with conscience-based objections have a duty to disclose this to their employer so they can accommodate patient access.

---

[3] Bob Roehr, "Assisted Dying in US and Canada: Controversy Subsides after Legalization," *The British Medical Journal* 360 (2018): k503.

[4] Thaddeus M. Pope, "Legal History of Medical Aid in Dying: Physician assisted Death in US Courts and Legislatures," *New Mexico Law Review* 48, no. 2 (2018): 267.

[5] American Medical Association, "Physician-Assisted Suicide," November 14, 2016, https://www.ama-assn.org/delivering-care/ethics/physician-assisted-suicide

[6] ANA Ethics Advisory Board, "ANA Position Statement: The Nurse's Role When a Patient Requests Medical Aid in Dying," *The Online Journal of Issues in Nursing* 24, no. 3 (2019).

[7] Ronit Y. Stahl, and Ezekiel J. Emanuel, "Physicians, not Conscripts-Conscientious Objection in Health Care," *The New England Journal of Medicine* 376, no. 14 (2017): 1380-1385.

Savulescu & Schuklenk make a similarly nuanced argument that speaks to the importance of the discussion.[8] While they do call for an end to conscience protection laws, they do so by demonstrating that objections are often better grounded in professional codes of ethics and scope of practice provisions. Their point, we believe, is not to silence professionals, but rather to protect patients by putting the focus on individual patient rights before a right to conscience-based objections by clinicians. Their approach puts the burden onto the clinicians, who have more power than a patient, by requiring them to base their objection in professional practice and code of ethics rather than the more ambiguous reference to "conscience." We believe that individual conscience-based objections are a legitimate and necessary moral exercise but that careful balance is needed to ensure that patient rights are not overridden by the same type of mistaken medical morality that led to faulty medical ethics by the Nationalist Socialists in Nazi Germany.

There are certainly bioethicists calling for removal of individual conscience objections, but not all.[9] The debate in bioethics is not one of consensus for blanket removal of conscience protections rather it is more nuanced and is over the balance between patient access to legal and professionally accepted care and respect for the individual conscience of their healthcare provider in order to avoid morally tenuous scenarios.

Assumption III

Fernandes & Ecret's conclusion that the Nazi regime indoctrinated physicians and nurses to act immorally provides an important and valuable lesson. It does speak to an element of individual conscience protections. However, we believe that discussion of hierarchy and moral silence by health professionals in the Nazi regime is more aptly analogized to the dangers of institutional conscience-based protections and institutional power, namely that a hierarchal structure such as a hospital can force its ideological beliefs on the community it serves, even coopting employees as morally apathetic agents of that hierarchy. Nazi physician's and nurse's individual objections were not made because they were willful participants in the regime's belief structure.

Institutional conscience protections allow healthcare institutions to decline provision of certain services based on their mission and values, even

---

[8] Julian Savulescu, and Udo Schuklenk, "Doctors have no Right to Refuse Medical Assistance in Dying, Abortion or Contraception," *Bioethics* 31, no. 3 (2017): 162-170.

[9] Mark R. Wicclair, "Is Conscientious Objection Incompatible with a Physician's Professional Obligations?" *Theoretical Medicine and Bioethics* 29, no. 3 (2008): 171-185.

if such therapies are considered medically indicated. Institutional conscience protections are widely supported by both federal and state laws.[10]

Fernandes & Ecret's arguments make the case for why institutions should not promote a particular "conscience" that might override individual conscience. This becomes particularly important, and relevant, when individual hospitals can promote conscience-based objections. Hospital conscience-based objections interfere with patients' rights to standard and legally authorized treatments similar to the way in which the Nazi regime reduced the rights of segments of the population, ensuring compliance from their clinicians.

I. Institutional Conscience

Individual conscience protections do not seem to have been the problem in the Holocaust, they could have spoken up, likely would have been punished. The absence of these laws does not seem to have been the main culprit in getting clinicians to act out their bidding. Rather, their analysis highlights the dangers being a part of a collective entity that encroaches on the rights of those they are supposed to serve in the name of an institutional belief structure, engaging in group think and moral silence. Fernandes & Ecret are correct that nurses and physicians have an obligation to speak up against institutional practices they conscientiously and morally object to.

Clinicians who many find themselves working in a, for instance, Catholic health care facility, are not constrained by objections to care, but rather they are constrained by the inability to provide care.[11] These are problems of institutional restrictions to care, the provider can voice their objection about a hospital not providing a legal medical service, but the patient still suffers a denial of rights, delays in care, and sometimes inadequate treatment. In fact, in some facilities a clinician cannot even counsel or advise a patient on how to access a service outside of the institution without a risk of employment violations.[12]

Fernandes & Ecret present an important and powerful analysis. But their analysis supports the conclusion that patient rights and protections are paramount. The ethical violations of Nazi physicians and nurses was not

---

[10] Nadia N. Sawicki, "Protections from Civil Liability in State Abortion Conscience Laws," *Journal of the American Medical Association* 322, no. 19 (2019): 1918-1920.

[11] Maryam Guiahi, "Catholic Health Care and Women's Health," *Obstetrics and Gynecology* 131, no. 3 (2018): 534-537.

[12] Debra B. Stulberg, Rebecca A. Jackson, and Lori R. Freedman, "Referrals for Services Prohibited in Catholic Health Care Facilities," *Perspectives on Sexual and Reproductive Health* 48, no. 3 (2016): 111-117.

a failure of conscience-based protections but a violation, by omission and commission, of people because of the "conscience" of the institution, or regime. They were complicit while the rights of those they were supposed to serve were reduced and eliminated. It is faulty to believe that one could have achieved justice through conscience protections in a wholly unjust society.

The reason this distinction is important, and why these lessons are relevant today is that conscience-based objections in healthcare have risen to the level of social discussion. However, there is a distinction that is getting lost in the discourse. This distinction is between the conscience of an individual clinician (nurse, physician) and the conscience of an entity such as a hospital (institutional conscience).

II. The Balance of Conscience and Rights

The balance between conscience and rights has been shifting in the US, as Catholic hospitals are merging with and acquiring hospitals around the country.[13] Problematic institutional consciences are not isolated to religiously-affiliated hospitals, this is our point. They are only currently emblematic of the dangers around allowing institutions to push particularly restrictive belief structures. One could imagine a secular institution employing some other restrictive conscience. As major policy initiatives are being attempted through the expansion of federal conscience rules that would expand what services and to whom services can be denied.[14] These policies favor institutional conscience over that of patient rights to access basic, standard, and legally accepted healthcare services, even when their own insurance policies allow for access. A clinician should maintain the right to refuse participation in controversial procedures, but not ones considered basic or urgent care. Institutional conscience policies also place physicians and nurses in a compromising position, one that can lead to moral distress in the face of concerns for employment violations.[15] This becomes the problem that Fernandes & Ecret warn of, institutional silencing of individual morals and professional codes of ethics leading to the harm of patients.

The shift towards institutional conscience protections also impacts the professional ethics of the clinical staff within the organization, just as Fernandes & Ecret detail in their account of Nazi physicians and nurses. While the authors

---

[13] Maryam Guiahi, "Religious Refusals to Long-Acting Reversible Contraceptives in Catholic Settings: A Call for Evidence," *American Journal of Obstetrics and Gynecology* 222, no. 4 (2019): S869.e1-S872.e5.

[14] Lawrence O. Gostin, "The 'Conscience' Rule: How will it affect Patients' Access to Health Services?" *Journal of the American Medical Association* 321, no. 22 (2019): 2152-2153.

[15] Guiahi, "Religious Refusals," S870.

make the argument that the lessons of the Shoah speak to the importance of individual conscience protections for physicians and nurses, we believe the more compelling and prescient argument is the protection of patients, and citizens, from institutional power. If hospitals have more power over patients, patients' rights may be violated, even making nurses and physicians complicit. The balance is recognizing the rights and responsibilities in both patients and clinicians without overly endowing power to the institution. The lessons from the Shoah, in Fernandes & Ecret's analysis, are that moral silence is unjust and institutional power needs to be checked. Individual clinician conscience regulations are too easily corrupted by institutional power. Without the dissolution of institutional conscience protections or the acceptance of conscientious provision protections, institutions will have too much power over patients and clinicians, and it is likely that history will sadly continue to repeat itself.

References

American Medical Association. "Physician-Assisted Suicide." November 16, 2020. https://www.ama-assn.org/delivering-care/ethics/physician-assisted-suicide.

ANA Ethics Advisory Board. "ANA Position Statement: The Nurse's Role When a Patient Requests Medical Aid in Dying." *The Online Journal of Issues in Nursing* 24, no. 3 (2019).

Bob Roehr. "Assisted Dying in US and Canada: Controversy Subsides after Legalisation." *British Medical Journal* 360 (2018): k503.

Fernandes, Ashley K., and DiAnn Ecret. "The Effect of Hierarchy on Moral Silence in Healthcare: What Can the Holocaust Teach Us?" *Conatus – Journal of Philosophy* 4, no. 2 (2019): 21-43.

Guiahi, Maryam. "Catholic Health Care and Women's Health." *Obstetrics and Gynecology* 131, no. 3 (2018): 534-537.

Guiahi, Maryam. "Religious Refusals to Long-Acting Reversible Contraceptives in Catholic Settings: A Call for Evidence." *American Journal of Obstetrics and Gynecology* 222, no. 4 (2019): S869.e1-S869.e5.

Gostin, Lawrence O. "The 'Conscience' Rule: How Will it Affect Patients' Access to Health Services?" *Journal of the American Medical Association* 321, no. 22 (2019): 2152-2153.

Pope, Thaddeus Mason. "Legal History of Medical Aid in Dying: Physician Assisted Death in US Courts and Legislatures." *New Mexico Law Review* 48, no. 2 (2018): 267-301.

Savulescu, Julian, and Udo Schuklenk. "Doctors Have No Right to Refuse Medical Assistance in Dying, Abortion or Contraception." *Bioethics* 31, no. 3 (2017): 162-170.

Sawicki, Nadia N. "Protections from Civil Liability in State Abortion Conscience Laws." *Journal of the American Medical Association* 322, no. 19 (2019): 1918-1920.

Stahl, Ronit Y., and Ezekiel J. Emanuel. "Physicians, Not Conscripts-Conscientious Objection in Health Care." *New England Journal of Medicine* 376, no. 14 (2017): 1380-1385.

Stulberg, Debra B., Rebecca A. Jackson, and Lori R. Freedman. "Referrals for Services Prohibited in Catholic Health Care Facilities." *Perspectives on Sexual and Reproductive Health* 48, no. 3 (2016): 111-117.

Wicclair, Mark R. "Is Conscientious Objection Incompatible with a Physician's Professional Obligations?" *Theoretical Medicine and Bioethics* 29, no. 3 (2008): 171-185.

Wolfe, Ian D., and Thaddeus M. Pope. "Hospital Mergers and Conscience-Based Objections – Growing Threats to Access and Quality of Care." *The New England Journal of Medicine* 382 (2020): 1388-1389.

interview

A. A. Long & D. Vertzagia · *Conatus* 5, no. 1 (2020): 111-122
DOI: http://dx.doi.org/10.12681/cjp.23324

Antiquity Revisited: A Discussion with Anthony Arthur Long

Anthony Arthur Long
University of California, Berkeley, USA
E-mail address: aalong@berkeley.edu
ORCID ID: https://orcid.org/0000-0002-7447-7237

Despina Vertzagia
National and Kapodistrian University of Athens, Greece
E-mail address: dvertzagia@philosophy.uoa.gr
ORCID ID: https://orcid.org/0000-0002-7894-8657

Abstract

A discussion on antiquity with Anthony A. Long, one of the most distinguished scholars in the field of ancient philosophy, would be engaging in any case. All the more so, since his two recently published works, Greek Models of Mind and Self (2015) and How to be Free: An Ancient Guide to the Stoic Life (2018), provide the opportunity to revisit key issues of ancient philosophy. The former is a lively and challenging work that starts with the Homeric notions of selfhood, and leads the reader all the way through classical and Hellenistic philosophical psychology; the latter is a profound analysis of the Stoic ethics that focuses in particular on its foundation and principles, followed by Long's re-worked translation of Epictetus' Encheiridion and carefully selected parts of his Discourses. Anthony Long kindly accepted the invitation to discuss several issues that are in the core of scholarly concern, sharing interpretations and thoughts that originate from his long acquaintance with the ancient literary tradition.

Key-words: *antiquity; Stoic ethics; Homer; Plato; Aristotle; soul; polis; philosophical psychology; freedom*

I. Homer revisited

Despina Vertzagia: Taking into consideration their diametrically opposed ontological and psychological background, to what extent would it be promising to compare Homer's and Plotinus' views on the immortality of the soul? For example, in the instance from *Odyssey*, where Plotinus claims that Hercules is portrayed as having "bifurcated identity," his "divine essence" doesn't seem to "focus on the life of the mind" inasmuch as Hercules "takes

delight in feasting with gods."[1, 2] Could you please explain which is the common place between Plotinus and Homer?

Anthony A. Long: The common point between Homer and Plotinus is simply the notion that a human soul (e.g. the soul of Hercules) can be in Hades after death as a shade, and simultaneously enjoy a divine immortal existence. However, Plotinus states at *Ennead I* [6.8] that Hercules was not a contemplative person; that is why only a part of him is with the gods. Plotinus clearly liked Homer's bifurcated Hercules, because he alludes to him again in *Ennead IV* [3.27]. There he uses the story of Odysseus's escape from Calypso, to illustrate his own recommendation to turn from physical beauty to spiritual beauty. Ancient philosophers often drew on Homer metaphorically or playfully to align him with their own system, as Plotinus does here.

Despina Vertzagia: If we compare Aristotle to Homer, could we assume that Aristotle partially restores the Homeric tradition due to the fact a) that his psychology does not recognize a human soul independent of the body, nor does he speak of immortality and post-mortem survival of the soul, b) that his perception of politics and his conception of courage remind us of the Homeric man? In your opinion how 'Homeric' could Aristotle taken to be?

Anthony A. Long: Aristotle reveres Homer as a great poet and likes to quote him. In the ways your question indicates, they do share a "psychosomatic" view of human identity to some extent, but I don't think Aristotle saw himself as going back to Homer, but rather as rejecting Plato's dualism. The Homeric psyche leaves the body at death for Hades, but it does survive as a ghost. There is nothing like that in Aristotle. For Aristotle the psyche is "the form of the body"[3] and the functioning of the living body. Homer never speaks of psyche in this way. His principal words for the mind are thumos and phrenes, which are not terms of Aristotelian psychology. As for politics and courage, I am not sure what Aristotle says that reminds you of Homeric man. Homer hardly talks about "politics," and I don't think he anticipates Aristotle's notion of courage as intermediate between rashness and timidity.

Despina Vertzagia: The memories or the words of the Homeric post-mortem "ghost"[4] aren't products of thinking or cognition? And if not, what

---

[1] Homer, *Odyssey*, 11, 601-602.

[2] Anthony A. Long, *Greek Models of Mind and Self* (Cambridge, MA & London: Harvard University Press, 2015), 19-20.

[3] Aristotle, *De anima*, 2, 1, 412a20.

[4] Long, *Greek Models of Mind and Self*, 30.

is their status, to the extent that their words actually have a meaning? And what about their instant "revival by blood?"[5] Could we assume that it stands for a kind of embodiment or resurrection? What is the precise status of the post-mortem Homeric soul?

Anthony A. Long: The question about the status of the "post-mortem Homeric ghost" is very interesting. It is called an eidolon, i.e. an unsubstantial form or phantom. It has a shadowy existence, but it is not alive. It can appear as a dream, and talk to the dreamer, as the shade of Patroclus appears to Achilles.[6] Odysseus revives the "souls" temporarily by giving them sacrificial blood to drink when they appear before him like vampire bats.[7] Based on their responses to him, they are frozen in time. They have a past but no future; so they are not "resurrected" or re-embodied. Their words have meaning but their speakers are said to be "mindless." Homer, however, is not a philosopher, and so his poems resist generalization or questions like "What is the precise status of these ghosts?" The shade of Teiresias is able to prophesy to Odysseus, and the hero converses with the shades of his mother, Agamemnon, and Achilles, as if they were temporarily with him as intelligent persons.[8] Their words are needed to advance the story, but like all the shades they are mere ghosts, imaginary rather than real.

Despina Vertzagia: Is there a concept of justice in the Homeric corpus? Leo Strauss claims that philosophy, natural right and justice emerged due to the questioning of the primeval identification of the good with ancestral.[9] As long as Homer is under the constellation of this notion, could we assume that there is a Homeric concept of justice?

Anthony A. Long: Yes. There is legal justice, requiring compensation for injury. It initiates Telemachus's anger with the aggressive suitors in the *Odyssey*. On the shield of Achilles in the *Iliad* a trial over homicide is depicted. Moral justice in Homer is what is socially approved or expected, with shame as its sanction. Homeric persons are expected to take responsibility for their actions, and when they do not, they suffer consequences in the form of human and divine disapproval. I don't think that Strauss's views are applicable to Homer.

---

[5] Ibid., 55.

[6] Homer, *Iliad*, 23, 60ff.

[7] Homer, *Odyssey*, 11, 146ff.

[8] Ibid., 11, 388ff, and 11, 467ff.

[9] Leo Strauss, *Natural Right and History* (Chicago & London: The University of Chicago Press, 1953), 86.

II. The dawn of the classical period

Despina Vertzagia: In the end of the chapter "Intimations of Immortality" you claim that "Heraclitus originated the ideal of the contemplative life."[10] Could you please elaborate on this?

Anthony A. Long: Heraclitus sought to interpret the Delphic maxim, "know yourself," by challenging his listeners to reflect on life and death, and to use their minds to investigate the nature of things, both the external world and ourselves as thinking beings. That is what I mean by saying that he originated the ideal of the contemplative life.

Despina Vertzagia: Soul-body dualism emerges in its clear form through the conflict between rhetoric and philosophy, as you claim in the chapter "Bodies, Souls and the Perils of Persuasion."[11] Could you provide some brief further justification for this claim?

Anthony A. Long: In Plato's early works, notably the *Apology*, *Gorgias*, and *Phaedo*, Socrates draws sharp contrasts between body and soul. The contrast had been made before, but its philosophical significance emerges with great clarity when we read these texts alongside Plato's account of the sophists Protagoras and Gorgias (who were experts in the art of persuasion) and Gorgias's *Encomium of Helen*. Gorgias had proposed that souls are unable to resist powerful rhetoric or the charms of beautiful bodies. On this view souls are naturally weaker than bodies. That is the reverse of the Socratic recommendation to subordinate body to soul. Thus Plato's dualism, with its elevation of rational soul over bodily desire, emerges as a major weapon in his focus on the perils of persuasion.

III. Plato and Aristotle on the soul and the polis

Despina Vertzagia: In the Preface of your *Greek Models of Mind and Self* you make reference to David Furley's advice that you should write a paper on Plato's psychology without even mentioning the word soul.[12] Could you elucidate on the reasons why you constantly refer to Furley's advice?

Anthony A. Long: The word "soul" has very different connotations in modern English from those of psyche in ancient Greek philosophy. Modern English speakers use the term "soul" chiefly to refer to the

[10] Long, *Greek Models of Mind and Self*, 85.

[11] Ibid., 88-124.

[12] Ibid., xiii.

emotional aspect of human nature (hence we speak of "soul" music) as distinct from rationality, or we may use "soul" to express the essence of something (saying, for instance, "she is the soul of discretion"). In Greek thought psyche refers primarily to the principle of biological life (hence Aristotle discusses psyche in his writings on nature), including the life of animals and even plants. In the case of animals psyche includes emotion, perception and desire, in the case of human beings it includes thought as well, and in all living creatures it is the source of metabolism and reproduction. By recommending us not to use the word "soul" in our essays on Plato, Professor Furley wanted his students to understand that Plato was exploring the foundations of human identity, not starting from an agreed sense of the meaning of psyche. When Plato writes about "immortality of psyche," it's OK to say "immortality of soul" because we do use the word "soul" to refer to a human being's spiritual identity. But psyche is also Plato's word for "mind," and we don't say "immortality of mind." There is no exact equivalent to ancient Greek psyche in modern science or philosophy of mind.

Despina Vertzagia: You argue that Aristotle is an exception to the philosophical background of his time to the extent that he doesn't perceive the soul as a thing assigned in "a definite location within the body," or within the universe.[13] a) Do you believe that in Plato's view the soul is extendible? And if not, to wit if Plato's soul is just independent and separable from the body, b) how would you interpret Aristotle's controversial reference to the active mind (nous poiêtikos) in *De anima* [III, 5]. Is this just a "Platonic slip" of Aristotle?

Anthony A. Long: According to Plato the psyche is an incorporeal substance. In the *Timaeus*, however, he gives the three parts of psyche bodily locations – brain, heart, and belly[14] – but psyche is not extendible in the sense that you could measure its size or divide it up. I don't think that Aristotle's "active intellect" is a "Platonic slip," as if Aristotle had forgotten his notion that the soul is the form of the body.[15] Just as actual objects outside the body are needed to activate perception, so (Aristotle reasons) an active (poiêtikos) intellect is needed to activate the embodied mind's capacity to think. Aristotle's philosophy of mind depends upon the notion that our embodied mind requires an external intellect that is already thinking, so as to actualize its own potentiality. My response is of course much too brief to do full justice to Aristotle's idea.

---

[13] Ibid., 24.

[14] Plato, *Timaeus*, 69e3-71b1.

[15] Aristotle, *De anima*, 2, 1, 412a20.

Despina Vertzagia: Unlike Aristotle, Plato gives priority to the intellectual rather than the ethical function of the soul.[16] Do you think that for Plato the intellectual is more important because it is a presupposition of the ethical, or both are perceived as completely identical functions of the soul?

Anthony A. Long: For both Plato and Aristotle, knowledge of changeless truths is the highest activity of the soul. The difference between them turns on the fact that, for Plato, such knowledge includes the spheres of ethics and politics, whereas for Aristotle these are branches of "practical" knowledge and not "pure science." For Plato, then, ethical understanding has a theoretical/mathematical basis (knowledge of the Forms). In Aristotle moral virtue depends on a combination of good habits and making the correct moral choices.

Despina Vertzagia: In the beginning of the chapter "The Politicized Soul and the Rule of Reason" you claim that Plato "was not committed to a single model of the mind/body relation and the structure of human identity."[17] Is it the tripartite theory of the soul that you have in mind? In your opinion, is it a different model that Plato provides in regard to the relation between the soul and the body, or just an elaboration of the 'mind-and-body-distinction' theory? And where could the body be 'located' in that context? Could we assume that it (the body) is identical with the appetitive part of psyche?

Anthony A. Long: As I explain at the end of my book's third chapter, Plato revised the extreme body/soul dualism of the *Phaedo*.[18] In that dialogue appetites are described as functions of the body. In the tripartite model (see *Republic*, *Phaedrus*, and *Timaeus*) appetites are functions of the soul. In this model, the soul (or the ego) is the agent of all desires; so you can't just say: "O, it was my body that made me do that greedy or lustful thing. No, it was I myself who did it because I wanted to."

Despina Vertzagia: You claim that Plato possibly "was not completely committed to the literal immortality of individual souls."[19] Is this doubt of yours based exclusively on *Timaeus* [90b-d], or is immortality of the soul in general a secondary issue for the Platonic psychology in your view? Is the theory of the immortality of the soul an enlarged allegory?

Anthony A. Long: In the *Symposium*, immortality is presented as the object of all mortal creatures' erotic desire to procreate. Being mortal, they cannot live forever, so they beget offspring as surrogates. There is no

---

[16] Long, *Greek Models of Mind and Self*, 121.

[17] Ibid., 125.

[18] Ibid., 122.

[19] Ibid., 153.

mention here of the soul's immortality. I think Plato would like to believe in that, but I don't think he proves it to his own satisfaction. Where he dwells on it at length in the *Phaedo*, its literal truth seems less important than the philosophical soul's affinity to the eternal Forms as distinct from the body and physical objects.

Despina Vertzagia: It is commonly argued by scholars that the polis to Plato is an enlarged soul. But couldn't it be the other way around? That is, is it possible that Plato conceived the structure of the psyche as in accordance to some patterns of the polis?[20]

Anthony A. Long: The *Republic's* analogy between the city and the soul has generated great scholarly discussion. I think the analogy works both ways. Socrates moves from the city's proposed three parts to the parts of the soul, but as I argue in my book, he politicizes the soul by distinguishing its proper ruler (reason) and its properly subordinate populace (appetites).[21] So, I agree that his account of the soul's structure presupposes these political concepts.

Despina Vertzagia: Is the *Republic* a political dialogue? Aristotle argues that Plato's political thought is despotic, because Plato discusses the polis by exactly the same terms he uses to discuss the oikos.[22] In your opinion, is Aristotle right in his view that Plato's political thought is despotic, or even authoritarian?

Anthony A. Long: Plato's political thought is unquestionably authoritarian, but authority in the *Republic* is based on reason's right to rule because only reason can organize the state so as to ensure the happiness of all. This sense of authority is not despotic in the modern sense of being simply based on power. Aristotle, however, thought that Plato was unduly committed to the unity and unification of the state as the condition of its stability.

Despina Verzagia: In the chapter "The Politicized Soul and the Rule of Reason" you claim that reason in Plato has desires.[23] Does this also imply that Plato understands all the parts of the soul in terms of desire, or would such an inference be challenged as explaining Plato through Stoic lenses? Furthermore, in the case we assume that reason is in fact a rational and hence a noble desire, how is it possible, as Plato says, that reason is also used "for wrongful purposes?"[24]

---

[20] Ibid., 150.

[21] See Chapter 4: "The Politicized Soul and the Rule of Reason," in Long, *Greek Models of Mind and Self*, 125-161.

[22] Aristotle, *Politics*, 1, 1252a8-18. See also Plato, *Statesman*, 258e-259c.

[23] Long, *Greek Models of Mind and Self*, 140-141.

[24] Ibid., 141.

Anthony A. Long: Plato uses the word epithumetikon to refer to the "appetitive" part of the soul that desires food, drink, etc. But he also, a bit confusingly, attributes desires to all parts of the soul: the rational part properly desires wisdom and the spirited part desires honour. So Plato does understand all parts of the soul in terms of desires. If people allow their appetites to govern their lives, they misuse reason by treating it instrumentally as the slave of their passions. As you rightly say, his giving desires to reason influenced the Stoic doctrine that passions are misjudgments by the rational faculty as distinct from irrational urges.

IV. Stoic Ethics

Despina Vertzagia: In both your *Greek Models of Mind and Self* and *How to be Free* you focus on the Stoic concept of assent (synkatathesis).[25] Do you think that this is a key concept for the psychology of the Stoics? What is its relation to the Socratic elenctic method?

Anthony A. Long: The mental function of assent (synkatathesis) is absolutely central to Stoic moral psychology. It signifies agency, decision, "going for something," commitment, determination. It has an affinity to what Aristotle calls "deliberative desire" (bouleutike orexis)[26] and also to what Cicero and Seneca call voluntas. It is the ancestor of our notion of the "will." Assent has a Socratic origin in the notion that our actions are always determined by what we accept as the best thing for us to do in our present circumstances. Thus Medea (the Stoics' favourite example of moral error) mistakenly assented to the thought that killing her children was the best thing for her to do, and therefore she acted accordingly.

Despina Verzagia: In your *Hellenistic Philosophy*, in the "Stoic Ethics" chapter in particular, you dedicate a few lines to the concept of opportune or timely behavior and action.[27] Could you provide an explanation of this notion in the context of Stoicism? To what extent does this notion explain the Stoic attitude? Is it inconsistent with the stoic teaching that we should accept all events in life as they are?

---

[25] Ibid., 188; Anthony A. Long, *How to be Free: An Ancient Guide to the Stoic Life, Epictetus, Encheiridion and Selections from Discources* (Princeton, NJ: Princeton University Press, 2018), 123-135; 115. See also Anthony A. Long, *Hellenstic Philosophy: Stoics, Epicureans, Sceptics* (Berkeley & Los Angeles: University of California Press, 1974), 91.

[26] Aristotle, *Nicomachean Ethics*, 1139a23.

[27] Long, *Hellenstic Philosophy*, 206-207.

Anthony A. Long: The Stoics used the word eukairos[28] to describe their wise man's "timely" character and behavior. Chrysippus defined the Stoic goal of life as "living according to experience of natural events."[29] Such experience ideally equips people to be excellent judges of what it is appropriate or opportune for them to do by assessing their external circumstances, abilities and social roles and functions (duties). At the limit, you might need to decide, whether it would be better to die rather than to live. Epictetus gives copious examples of such "timely" behaviour. The essence of ancient Stoicism was not passively "accepting events as they are" (that is a modern distortion), but making best possible use of events: as Epictetus said with reference to Socrates, he always played the ball well, even in prison.[30]

Despina Vertzagia: Since Stoic moral theory focuses on the integrity of the individual, would we be justified to consider it a self-centered ethics? How is the Stoic man related to the community and to the other? Is there any notion of altruism in the stoic moral universe, or is care for the family and the community just a reflection of instinctive self-preservation, as it is usually argued?

Anthony A. Long: All ancient ethical theories are "self-centered" in the sense that they recommend how to achieve eudaimonia, the best possible life for oneself. They are no less socially oriented because they all treat ethical excellence (arete) – courage, justice, etc – as either the most important ingredient of eudaimonia (Aristotle and Plato), or an essential instrument of eudaimonia (Epicurus), or entirely identical to eudaimonia (Stoicism). According to Stoicism, human beings are born with instincts both for self-preservation and for family and community life. In caring for other people, the ideal Stoic is also caring for herself, i.e. her own excellence as a virtuous person. She is motivated by the desire to activate her virtues because they are the basis of her living well and successfully. She is not altruistic in the sense that she acts for the sake of others instead of herself, or by sacrificing her own interest. In doing good to others and desiring so to act, she is simultaneously desiring and doing good for herself. There is of course, much more to say on this very big subject. Your readers might care to read what I have written in Chapter 7 of my book, *Epictetus: A Stoic and Socratic Guide to Life*.[31]

---

[28] *SVF*, 3, 521.

[29] See for example *SVF*, 3, 5-6; 9-10.

[30] Epictetus, *Discourses*, 1, 12.

[31] Anthony A. Long, *Epictetus: A Stoic and Socratic Guide to Life* (Oxford: Oxford University Press, 2002), 180-206.

Despina Vertzagia: Health for the Stoics, and especially for Epictetus, is classified as a 'preferred indifferent' (adiaphoron), since it is not "up to us."[32] But it is a common ground that our health depends heavily on our life-style and nutrition choices and habits. To the extent that health is up to us, is it still morally indifferent?

Anthony A. Long: When the Stoics say that health is "a preferred indifferent" they mean that while we naturally prefer health to sickness, good health is not essential to living a good life: you can be a good and happy person even if you are a permanent invalid, and you can be a bad and unhappy person even if you have excellent health. The Stoics acknowledge the obvious fact that our state of health greatly depends on our life-style. Not only that, because good health is naturally preferable to sickness, it is incumbent on us to do everything possible to live a healthy life; otherwise we could not be living in agreement with our natural instincts. Trying to be healthy is morally correct for a Stoic. But the achievement of good health is morally indifferent because it depends on many things that are beyond our control.

Despina Vertzagia: What is the relation between "the Stoic model of mind and our own phenomenological experience,"[33] the comparison by which you conclude the last chapter of *Greek Models of Mind and Self*? How applicable is the Stoic model of life to a "disenchanted" world?

Anthony A. Long: The Stoic focus on the mind's potential autonomy and inner freedom can be of great value today as we attempt to navigate the complexities of modern life. This is not simply my opinion. In the few years since I wrote *Greek Models of Mind and Self*, ancient Stoic writers, especially Epictetus and Marcus Aurelius, have been enjoying an extraordinary renaissance as guides to life. Their teaching is incorporated in works on mindfulness, the search for meaning, self-help, and dealing with anxiety and other forms of mental trauma. What I myself want to emphasize most strongly, is Stoicism's focus on human dignity and the mind's emotional resourcefulness. Modern life imposes great strains on people, especially young people who are dealing with the difficulties of employment, personal relationships, and pressures on physical and mental health. While all too many people are victims of circumstances way beyond their control, Stoicism proves itself capable even in such cases of providing solace and inner strength. It challenges people to discover their privileged identity as the only animal with the capacity for self-determination.

---

[32] Long, "Introduction," in Long, *How to be Free*, xxvii-xxxv.

[33] Long, *Greek Models of Mind and Self*, 195.

References

Aristotle. *Nicomachean Ethics*. Translated by H. Rackham. Loeb Classical Library 73. Cambridge, MA: Harvard University Press, 1932.

Aristotle. *Politics*. Translated by H. Rackham. Loeb Classical Library 264. Cambridge, MA: Harvard University Press, 1932.

Aristotle. *On the Soul. Parva Naturalia. On Breath*. Translated by W. S. Hett. Loeb Classical Library 288. Cambridge, MA: Harvard University Press, 1957.

Homer. *Iliad*. Volume II, Books 13-24. Translated by A. T. Murray, and W. Wyatt. Loeb Classical Library 171. Cambridge, MA: Harvard University Press, 1925.

Homer. *Odyssey*. Volume I, Books 1-12. Translated by A. T. Murray, and G. E. Dimock. Loeb Classical Library 104. Cambridge, MA: Harvard University Press, 1995.

Long, Anthony A. *Epictetus: A Stoic and Socratic Guide to Life*. Oxford: Oxford University Press, 2002.

Long, Anthony A. *Greek Models of Mind and Self*. Cambridge, MA & London: Harvard University Press, 2015.

Long, Anthony A. *Hellenstic Philosophy: Stoics, Epicureans, Sceptics*. Berkeley & Los Angeles: University of California Press, 1974.

Long, Anthony A. *How to be Free: An Ancient Guide to the Stoic Life, Epictetus, Encheiridion and Selections from Discources*. Princeton, NJ: Princeton University Press, 2018.

Plato. *Statesman. Philebus. Ion*. Translated by H. N. Fowler, and W. R. M. Lamb. Loeb Classical Library 164. Cambridge, MA: Harvard University Press, 1925.

Plato. *Timaeus. Critias. Cleitophon. Menexenus. Epistles*. Translated by R. G. Bury. Loeb Classical Library 234. Cambridge, MA: Harvard University Press, 1929.

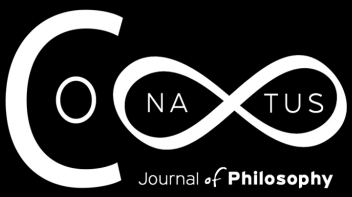
Plotinus. *Ennead, Volume I: Porphyry on the Life of Plotinus. Ennead I*. Translated by A. H. Armstrong. Loeb Classical Library 440. Cambridge, MA: Harvard University Press, 1969.

Plotinus. *Ennead, Volume IV*. Translated by A. H. Armstrong. Loeb Classical Library 443. Cambridge, MA: Harvard University Press, 1984.

Strauss, Leo. *Natural Right and History*. Chicago & London: The University of Chicago Press, 1953.

Stoicorum Veterum Fragmenta. Volume 3. Edited by Hans von Arnim. Eugene, OR: Wipf and Stock, 2016.

CONATUS 5, NO. 1 WAS DESIGNED BY ACHILLES KLISSOURAS, TYPESETTED IN PF AND UB TYPEFACE, PRODUCED BY THE NKUA APPLIED PHILOSOPHY RESEARCH LABORATORY AND PRINTED IN 100 GR. ACID FREE PAPER. THE ELECTRONIC VERSION OF THIS ISSUE IS HOSTED BY THE NATIONAL DOCUMENTATION CENTER.



www.conatus.philosophy.uoa.gr
p-ISSN: 2653-9373
e-ISSN: 2459-3842