Conf. Cephalal. et Neurol. 2022; Vol. 32, N. 1: e2022007 

Multidisciplinary Research In Neurosciences

# Dostoevsky's Karamazov Brothers and the problem of biological determinism

*Sofia E. Walters*
Scuola Normale Superiore, Pisa - E-mail: sofia.walters@sns.it

**Abstract.** In this article I show how Dostoevsky, in the novel "The Brothers Karamazov", develops a model of how the mind-body problem and the problem of neurological determinism interact with the notion of personal agency and free choice. I use current conceptual frameworks to untangle the conceptual knots of this analysis, such as the layering of biological determinism; the relationship between reductionism, antireductionism and cognition; the modern analysis of free will and agency; the role of emotions within ecological contexts.

**Key words:** Dostoevsky, Free will, Neurological determinism, emotion and ecological cognition

## Introduction

Between 1879 and 1880 Fyodor Dostoevsky published his last novel, "The Karamazov Brothers". The novel deals with the relationships between a morally derelict father, Fyodor Pavlovich Karamazov, his three legitimate sons - Dmitri, Ivan and Alëša - and his illegitimate son Smerdjakov. The plot of the novel follows the interactions between these characters as they deteriorate and become increasingly complex, following the assassination of Fyodor, the suspicions surrounding this murder and the subsequent incarceration of Dmitri, the eldest son.

This plot allows Doestoevsky to develop multiple themes, including those of imprisonment, disease and death. In this article I will try to focus on that of determinism, which I believe to be strictly linked to all the above mentioned themes and to Dostoevsky's own conception of the mind-body problem. The inclusion and the analysis of these themes stems from the development of the main characters, each of which, despite their realism and complexity, chiefly embodies a specific character trait[1]: Fyodor is dissolute, Dmitri is passionate, Ivan is rational and Alëša is spiritual. Smerdjakov apparently resists a classification of this kind, and is simply considered a bitter and diseased man[2].

These characters show how the different sides of human nature (passion, intellect, spirituality and determinism itself) relate to the problem of human agency when confronted with the ramifications of the scientific results of the time. I will focus mainly on the roles of passion and determinism.

## Modern biological determinism

One central theme in Dostoevsky's work that I want to focus on is biological determinism. Currently, determinism is defined as follows:

1 - This, of course, is a simplification and does not account for the changes in the characters' personalities within the novel itself.

2 - Despite this remark, I will try to briefly show how Smerdjakov's character is intrinsically linked with the problem of biological determinism.

> The world is governed by (or is under the sway of) determinism if and only if, given a specified way things are at a time t, the way things go thereafter is fixed as a matter of natural law.[3]

Biological determinism, then, is the idea that biological beings necessarily follow and are shaped according to biological laws. The first thing that needs to be noted, then, is that to understand biological determinism it is first essential to understand *which laws* we are referring to. This question seems odd and irrelevant, but it helps us focus on how the theme of biological determinism can become philosophically relevant. The laws of physics are obviously the first thing that comes to mind, since biology depends on our understanding of physics. However, in describing biological processes it is customary to view them on a different complexity level. Still, it is not clear exactly which level this might be: are biological entities determined on a purely molecular level? On a physiological level? Or rather on an ethological or evolutionary level? To make sense of this kind of process it is fundamental to consider all of these levels at once, as highlighted by R. Sapolsky in his book "Behave - The biology of humans at our best and worst"[4]. Sapolsky divides the problem into different steps: to explain what causes a specific behaviour one has to first study what has happened in the body of the agent a second before the behaviour takes place, then what happened seconds to minutes before, going further and further backwards in time. This transforms the analysis of the behaviour in the solution of the following questions:

1. What neural stimulus has caused the behaviour?
2. What hormones have caused the biological entity to be in the state it was in when the stimulus occurred?
3. What genetic predisposition was there in this individual for that hormonal set-up?
4. ….

*N.* Finally, what ecological pressures or advantages lead to the development of that particular genetic make-up?

The necessity of such a complex analysis stems from the fact that each level depends on the others, and answering only one of them would leave us still begging the question on the phenomenon. Of course, an analysis of this kind presupposes determinism itself: while it is quite easy to argue that a particular phenomenon is undetermined by looking at the phenomenon from a temporally restricted point of view, it becomes increasingly difficult as the scope of the inquiry widens. A further consequence of this layering of analysis is the fact that the debate on whether phenomena such as behaviour are primarily influenced by nature or nurture ceases to make sense: ultimately both internal and external factors play instrumental roles in affecting the behaviour of biological beings, and need to be considered together in an ecological perspective. It is not possible to give models of the functioning of biological systems without taking into account their evolutionary and contextual salience. While this is particularly clear within the background of modern hard sciences, this approach has not necessarily been fully and explicitly employed within the research in cognitive science and philosophy, where abstract models, sometimes hardly ecologically relevant, still find fertile ground. This distinction rests on the fact that, while there is absolutely no reluctance to accept that processes such as digestion or perception might be biologically determined, the same does not hold true for other kinds of phenomena, that we inherently think of as deliberate and in our power - e.g. thought, choice and action. Processes of the former kind have traditionally been under the scope of hard sciences, whereas processes of the latter have been primarily investigated by soft sciences, up to recently.

Of course, the distinction between these processes is never completely clear cut (except when referring to involuntary reflexes). For instance, when speaking about the determinism of perception it has become clear that there are plenty of cases where perception

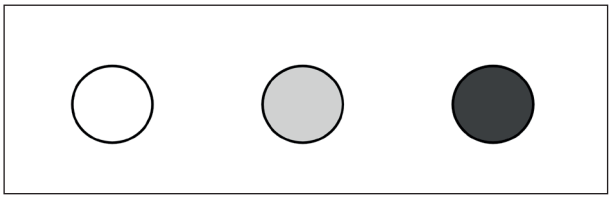

**Figure 1**.

3 - Hoefer, Carl, "Causal Determinism", The Stanford Encyclopedia of Philosophy (Spring 2016 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/spr2016/entries/determinism-causal/>.

4 - Sapolsky, Robert M.. "Behave: The Biology of Humans At Our Best and Worst". New York, New York: Penguin Press, 2017. Print.

itself can be manipulated and in which it has been found to be culturally bound. Such is the case of ambiguous representations, which allow more than one interpretation. An extremely famous example is that of the rabbit-duck illusion discussed in L. Wittgenstein's "Philosophical Investigations"[5]. A more relevant example, however, is that mentioned by P. Feyerabends "Dialogue on the method"[6]. The author cites the case of the image displayed in Figure 1.

When asked if the three elements in the picture look alike, different reasoners give different answers, which are statistically culturally determined. For example, western reasoners tend to answer positively because they interpret the three elements as three circles, albeit with different colours. When asked, on the other hand, Uzbek farmers answered to the contrary. In their view the three elements represented a bracelet, the moon and a coin respectively. As such, the Uzbek farmers did not feel that there was a strong similarity between them. An experiment of this kind shows how it is in principle possible to shape how we perceive some stimuli by working on our automatic representations, thus deliberately choosing what to believe. However there are at least two important objections to this line of argumentation: first of all, the fact that it is possible to have deliberate control on our beliefs does not imply that that is our usual epistemic stance (and it is clear that it is not: we do not usually have conscious and deliberate control on our beliefs); secondly, a sceptic would simply argue that our choice would be itself determined by biological and social factors we are not consciously aware of, thus making us fall back into determinism. It seems, then, that perception will fall into the category of "determined process", but I argue that the reason this happens is not simply because we find these objections decisive. Remarks such as these are usually accepted when directed to perception, but are not when referring to other processes which we consider to be more directly *mental*. For example, noting that we commonly make decisions unconsciously, both from a Freudian point of view and from a heuristic perspective[7], has not sealed the deal on the notion that our choices are out of our control. From a sceptical point of view this observation would be equivalent to saying that, while sometimes we perceive our choices as deliberate, since it has been experimentally proven that they most often are not, the less costly assumption should be that they simply are not. The fact that this argument is not readily accepted is a symptom of researchers' unwillingness to let go of certain theoretical assumptions that thrive within human sciences, such as those of agency, responsibility and free will. In other words: we consider the mental within our control not because there is some intrinsic difference between these processes and obviously deterministic ones, but rather because we have more philosophical assumptions on it. The gradual encroachment of hard sciences on soft ones, however, has progressively limited the importance of these assumptions, and it is quite clear that this process initially flourished in the twentieth century. Dostoevsky struggled with this process. To better understand what theoretical notions are actually part of the concept of biological determinism and which are mainly hidden assumptions, then, it can be useful to untangle the problem starting with Dostoevsky's outlook on the theme.

## Dostoevsky's mind-body problem

There are multiple instances where the plot of "The brothers Karamazov" forces the reader to focus on Dostoevsky's way of interpreting the mind-body problem[8]. The MB problem is the theoretical problem of describing how mind and body interact and deciding whether they can be considered one and the same. Dostoevsky is concerned with this problem because, while we would intrinsically place our thought (and, thus, our mind) within the realm of phenomena that are in some ways up to us, the encroachment of hard sciences had already started to change this and Dostoevsky was very well aware of this tendency.

In particular, there are two main points where

5 - Wittgenstein, Ludwig, and G E. M. Anscombe. "Philosophical Investigations". Oxford, UK: Blackwell, 1997. Print.

6 - Feyerabend, Paul. "Dialogo sul Metodo". Translation by R. Corvi, Edizioni Laterza, 2007.

7 - Kahneman, Daniel. "Thinking, Fast and Slow". New York: Farrar, Straus and Giroux, 2011.

8 - Henceforth, MB problem.

Dostoevsky shows his awareness of the problem and of its religious, ethical and practical ramifications:

1. when he describes the funeral rites of the Starec Zosima (Book VII, Chapter 1: *The odor of corruption*);
2. when Alëša and Dmitri discuss while the latter is in prison (Book XI, Chapter 4: *A hymn and a secret*).

*The odor of corruption*

In this instance the MB problem is focused on the body as a whole. Dostoevsky conveys his perplexities on the relationship between mind and body by narrating about the smell of the body of the deceased Starec Zosima, the spiritual guide of the youngest Karamazov brother, Alëša. Dostoevsky plays on the fact that in the religious conception of the time there are expectations about how the body of Zosima, a character who is considered a living saint, is supposed to appear after his death. Specifically, it is extremely surprising and disappointing for the people who believed in Zosima's sanctity that his body begins to decay and smell very quickly after his death, as show the following lines:

> When, still before dawn, the body of the elder, prepared for burial, was placed in the coffin and carried out to the front room, the former reception room, a question arose among those attending the coffin: should they open the windows in the room? But this question, uttered cursorily and casually by someone, went unanswered and almost unnoticed unless it was noticed, and even then privately, by some of those present, only in the sense that to expect corruption and the odor of corruption from the body of such a deceased was a perfect absurdity, even deserving of pity (if not laughter) with regard to the thoughtlessness and little faith of the one who had uttered the question. For quite the opposite was expected. Then, shortly after noon, something began that was first noticed by those coming in and going out only silently and within themselves, and even with an apparent fear of communicating the thought that was beginning to form in them, but which by three o'clock in the afternoon had manifested itself so clearly and undeniably that news of it spread instantly all over the hermitage and among all the pilgrims visiting the hermitage, at once penetrated the monastery as well and threw all the monks into consternation, and, finally, in a very short time, reached town and stirred up everyone there, both believers and unbelievers. [...] The thing was that little by little, but more and more noticeably, an odor of corruption had begun to issue from the coffin, which by three o'clock in the afternoon was all too clearly evident and kept gradually increasing.[9]

Dostoevsky here combines two sides of the problem of determinism (specifically the MB problem and moral responsibility) into one argument, which creates scandal in the eyes of the spectators

1. the body of the deceased is intrinsically connected to his soul,
2. there are bodily characteristics which are signs of the state of the soul (i.e. the soul influences the body),
3. the foul smell of the corruption of the body is an indication of the moral corruption of the soul.

What is particularly interesting is that Zosima's funeral is important both for the believers and for the non-believers. The non-believers reject the whole argument because they reject the idea that there is such a thing as a soul, while perfectly content with the notion of a body. The believers, on the other hand, have two possible ways of maintaining their beliefs while watching such a spectacle: they can either reject (1.) or (2.). Rejecting (1.) would mean essentially asserting the primacy of the soul on the body (which thus becomes a sort of platonic prison of the soul itself); rejecting (2.) would give way to a complex understanding of the relationship between body and soul. Interestingly, and quite unexpectedly, the believers seem to choose the second option ("*there was no Orthodox dogma that the bodies of righteous men are necessarily incorruptible, it was only an opinion*"). In other words: the fact that there is an incredibly tight relationship between mind and body does not intuitively represent a problem to the religiosity of the believers of Dostoevsky's time. What is problematic is assuming that from the state of the body it is possible to infer something about the connected soul, but there is no great scandal in accepting (1.). This highlights an important development in the

9 - Dostoyevsky, Fyodor. "The Brothers Karamazov". Translation by Richard Pevear and Larissa Volokhonsky, 1992, pp. 279-280.

understanding of biological determinism in Dostoevky's time: there was no perception of a deterministic problem when considering the body as a whole.

Although the modern understanding of determinism should prepare us to see no difference between the determinism of the body in its entirety and the determinism of sections of it, it is clear that such a distinction existed in the perspective of the time. This is particularly highlighted by the fact that, even though he judges it necessary to dedicate a whole chapter on the odour of Zosima's body and its religious implications, the author's focus is not on the body itself, but rather on the fact that the religiosity of the people who are present is full of superstition. The problem introduced, therefore, stems from the MB dualism and its repercussions, but does not hinge on it: having a determined body is not scandalous as much as believing it to be a product of moral responsibility.

*A hymn and a secret*

In this chapter Dmitri, the eldest and the most passionate of the brothers, is held in jail accused of the murder of his father. The real murderer, however, is his half-brother Smerdjakov, who has managed to frame Dmitri by giving himself an aliby for the night of the murder with a fake epilectic crisis. Dmitri, while in jail, meets various other characters, including a cynical and intellectual friend, Rakitin[10], and his spiritual brother, Alëša. Dostoevsky introduces the theme of biological determinism by making Dmitri and Alëša discuss a conversation Dmitri previously had with Rakitin about the new discoveries of science.

> "[...] The fact is ... on the whole ... I'm sorry for God, that's why!"
>
> "What do you mean, sorry for God?"
>
> "Imagine: it's all there in the nerves, in the head, there are these nerves in the brain (devil take them!) ... there are little sorts of tails, these nerves have little tails, well, and when they start trembling there... that is, you see, I look at something with my eyes, like this, and they start trembling, these little tails… and when they tremble, an image appears, not at once, but in a moment, it takes a second, and then a certain moment appears, as it were, that is, not a moment - devil take the moment - but an image, that is, an object or an event, well, devil take it - and that's why I contemplate, and then think... because of the little tails, and not at all because I have a soul or am some sort of image and likeness, that's all foolishness. Mikhail explained it to me, brother, just yesterday, and it was as if I got burnt. It's magnificent, Alyosha, this science! The new man will come, I quite understand that… And yet, I'm sorry for God!"
>
> "Well, that's good enough," said Alyosha.
>
> "That I'm sorry for God? Chemistry, brother, chemistry! Move over a little, Your Reverence, there's no help for it, chemistry's coming! And Rakitin doesn't like God, oof, how he doesn't! That's the sore spot in all of them! But they conceal it. They lie. They pretend. [...]"[11]

From this extract it is clear that, in focusing on biological determinism, Dostoevsky feels the need to hone in on the notion of neurological determinism specifically. While the body did not represent an obstacle to his understanding of human behaviour, it is clear that the brain and its functioning do. Just as well, in a single paragraph Dostoevsky presents the problem (I will refer to it as the problem of the "trembling tails") and the author's solution to it, expressed by the sentence "*And yet, I'm sorry for God!*". Before disentangling the conceptual knots that can be derived from these few lines, it is useful to give a more complete overview of the fundamental themes that accompany that of neurological determinism in this novel.

- **The tale of the prisoner:** the first salient observation on the subject is that the whole theme of neurological determinism is introduced by a prisoner's speech. Specifically, not merely by a generic prisoner, but by an innocent prisoner awaiting trial and fully convinced that he is going to be found guilty. This puts the whole theme into perspective. Dmitri, in jail, having little time and feeling the urgency to discuss important matters with his brother (who he considers to be the only one he can fully trust), spends most of it talking about the problem

10 - It is also interesting to point out that Rakitin, who is a relatively minor character, has a prominent role in this chapter, but had been explicitly mentioned in the previously cited chapter as well.

11 - Id. p. 499.

of determinism and how God can fit into a world of this kind. A reader, then, understands both how high the stakes are on this conceptual problem and its possible ethical implications: with an extremely effective semantic move Dostoevsky tells the reader that were one to find herself hostage to the problem of neurological determinism without finding a way out, she would become prisoner of it through no fault of her own. The cage of neurological determinism echoes the cage Dmitri is in and emphasises the fact that he has no responsibility for being in that situation: as much as he had no say in the murder of his father, if one accepted neurological determinism one would have to admit Dmitri had no say in anything whatsoever.

- **The diseased culprit:** Dostoevsky struggled for two thirds of his life with epilepsy[12], which had an extremely negative impact on his quality of life, but was intrinsically connected to his literary endeavours. Dostoevsky, in his correspondence, got to the point of confessing:

> The thing is that, for twenty-five years now, I have been suffering from epilepsy, which I contracted in Siberia. This illness has gradually deprived me of the ability to remember faces and events, to such an extent that I have (literally) even forgotten all the themes and details of my novels and, since some of them have never been reprinted since they were first published, I actually have no idea of what they are about.[13]

Besides his first-hand experience with epilepsy, Dostoevsky was also familiar with the medical literature on the subject and on neural mechanisms. A trace of his expertise can be found in the novel, when he cites Claude Bernard[14], a famous neurologist of his time. Bernard was well known for his opinions about the relationship between brain chemistry and mental processes. He is cited as stating:

> The philosophers whom we speak about accept the idea that the inferior phenomena of animals can depend on determinism… but not the superior psychological ones. Thus, in man we distinguish vital processes which lie within determinism and others that are out of it. In our opinion, physiological determinism can have no exception. [...] This is an absolute determinism: it expresses the fact that the psychological dimension cannot exist without the physical and chemical world.[15]

All this practical and theoretical understanding of brain functions and epilepsy influenced Dostoevsky, leading him to describe it in his writings, where a good deal of fundamental characters suffer from it. Specifically, in The Brothers Karamazov, Ivan, the intellectual brother, is said to suffer from some sort of brain fever and Smerdjakov, the culprit of the father's murder, is an epilectic. It is important to point out that Smerdjakov is an epilectic because of his ambiguous and alien status within the plot of the novel[16]. Smerdjakov is the fourth Karamazov brother, but he is never really cited as such; as already mentioned, he does not embody one side of the human soul as do the others; finally, he is raised as a servant, abused and neglected in his infancy. Epilepsy emerged as a consequence of this abuse: after being stricken by the man who raised him, Smerdjakov hit his head and developed the disease. At the same time, he lost his ability to forgive the abusers for the abuse committed towards him, thus becoming increasingly bitter. Dostoevsky, thus, creates a character where the ability to choose his course of action disappears and is somewhat substituted by the presence of a disease that not only has external symptoms, but shapes his whole thought process. This is particularly clear from the fact that, according to Dostoevsky himself, epilepsy caused cognitive

12 - Maze, J. R. "Dostoyevsky: Epilepsy, Mysticism, and Homosexuality". *American Imago*, vol. 38, no. 2, The Johns Hopkins University Press, 1981, pp. 155–83, http://www.jstor.org/stable/26303748. And, Iniesta, Ivan. "Epilepsy in Dostoevsky", Editor(s): Anne Stiles, Stanley Finger, François Boller, Progress in Brain Research, Elsevier, Volume 205, 2013, Pp. 277-293.

13 - In Iniesta, Ivan. "Epilepsy in Dostoevsky", Editor(s): Anne Stiles, Stanley Finger, François Boller, Progress in Brain Research, Elsevier, Volume 205, 2013, p. 291.

14 - Dostoyevsky, Fyodor. "The Brothers Karamazov". Translation by Richard Pevear and Larissa Volokhonsky, 1992, p. 498.

15 - Bernard, Claude. "Leçons sur les Phénomènes de la Vie", in Conti, F. "Claude Bernard's Des Fonctions du Cerveau: an ante litteram manifesto of the neurosciences?", in Nature Reviews of Neuroscience, 2002. My translation.

16 - For an interesting perspective on this character see Cohen, Sharon. "'Balaam's Ass': Smerdyakov as a Paradoxical Redeemer in Dostoevsky's The Brothers Karamazov." Christianity & Literature, vol. 64, no. 1, Dec. 2014, pp. 43–64, doi:10.1177/0148333114552772.

consequences such as long periods of depression[17]. Of course, the loss of personal agency and the emergence of a cognitive dysfunction are not causally connected, but it is interesting to point out that they both coexist within the same character representing his most obvious personality traits, thus instilling in the reader the doubt that the abused brother might, after all, be to some extent not responsible for his own actions.

- **Sorrow and freedom:** it is important to remark that the discussion with Rakitin does not have the effects that we would intuitively expect in Dmitri. Dmitri, as said, is an incredibly headstrong and passionate character, who follows what he believes in without hesitation. Rakitin manages to convince Dmitri completely from an intellectual standpoint of what he says, but Dmitri does not follow this line of thought where we might believe. He accepts that "the new man is coming" and the philosophical implications of neurological determinism, but that does not lead him to reject God or his beliefs on the existence of a soul. His response to the problem of determinism is one of sorrow regarding the abandonment of God: he understands the emergence and inescapability of nihilism that follows the acceptance of complete materialism, but is sad because of it. This sadness leads him in an unexpected direction: he has a spiritual awakening and finds himself finally free, even when unjustly convicted. He then raises an hymn to the freedom of his soul, despite what he has just learned about the determinism of his thoughts.

## Disentangling the problem of the trembling tails

In trying to make sense of the philosophical implications of Dostoevsky's account of the problem of biological / neurological determinism, it is necessary to analyse a few conceptual subtleties in his approach which can be better understood from a modern perspective. Specifically:

1. What does it mean to accept neurological determinism from the point of view of the MB problem in Dostoevsky's approach?
2. What exactly is the scope of the problem of neural determinism in The Brothers Karamazov? What are its ethical implications?
3. Is Dostoesky's solution effectively viable?

### *The notion of reductionism*

A phenomenon A is *reducible*[18] to a phenomenon B if and only if A is completely made of (or explainable by) the set of elements of B and their interactions. A first case of a reduction could be exemplified by how a simplified desert is reducible to a set of grains of sand arranged in a certain order and occupying a certain space: this practically means that the desert is *nothing more* than the grains of sand themselves that compose it.

This reduction, as can be seen in the abovementioned definition, can be of at least two kinds: ontological and epistemological. I will give two simplified examples:

1. A desert is made up of the grains of sand that compose it, arranged in a certain order and occupying a certain space, and of their interactions.
2. A kingdom is made up of all the people who compose it, arranged in a certain order and occupying a certain space, and of their interactions.

Intuitively, there is a difference between the two examples, in that a kingdom might just as well be made of those elements and their interactions, but this tells us little about it; whereas the same is not true for the desert. Namely, a desert is nothing more than a collection of grains of sand arranged in a specific way (which, of course, needs to be described in order for a reasoner to have understanding of the phenomenon "desert"). However, the explanatory efficacy of the definition in (2.) is much weaker: to properly understand what a kingdom is a reasoner would need a wider net of information than the simple list of the people who compose it and their interactions. To understand what a kingdom is, for instance, one might need to know concepts such as those of "institution", "politics", "social exchange", which do not figure within a bare list of the components of a kingdom even in our best quantitative

17 - Iniesta, Ivan. "Epilepsy in Dostoevsky", Editor(s): Anne Stiles, Stanley Finger, François Boller, Progress in Brain Research, Elsevier, Volume 205, 2013, Pp. 279.

18 - Brigandt, Ingo and Alan Love, "Reductionism in Biology", *The Stanford Encyclopedia of Philosophy* (Spring 2017 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/spr2017/entries/reduction-biology/>.

models[19]. In other words, there is a *qualitative*, conceptual and theoretical dimension to what a kingdom is that cannot be found at the level of analysis that would be required to properly speak of the reduction of A (the kingdom) to B (its elements and their interactions). Nonetheless, it might still be that a kingdom is entirely reducible to its elements and their interactions. This example shows that there is a gap to be closed between two ways of understanding reducibility.

Specifically, a phenomenon can be reducible to another from an *ontological* point of view, from an *epistemological* point of view or from both (and maybe also from neither). An ontological reduction happens when there *is* nothing more to the phenomenon A than what is in the phenomenon B. This is clearly the case of the desert, which is nothing more than its parts. Instead, we have an epistemological reduction when there is no explanatory gap between the mentioned levels. This, again, is clearly the case of the desert, where no further explanation is needed other than its composition. A kingdom, on the other hand, might be ontologically reducible to its components, and might also be not reducible from an epistemological point of view with our current models of it.

This analysis creates four possible cases of non-reducibility:

1. Absolute ontological non reducibility (AON): the phenomenon A is *intrinsically something more* than the pure collection of its components and their interactions.
2. Current ontological non reducibility (CON): our best current quantitative models of the phenomenon A do not show how A can be reduced to the level B.
3. Absolute epistemological non reducibility (AEN): the phenomenon A is *intrinsically unexplainable* on the level B.
4. Current epistemological non reducibility (CEN): our best current (quantitative) models of the phenomenon A do not show how it can be explained on the level B.

These four cases are plausible ways to understand various positions on the subject, but not all of them can necessarily be held or have been held by researchers. There are two things that are interesting to notice. First of all, when referring to current non reducibility one automatically takes an epistemological standpoint: it shifts the focus from what is intrinsically so and so to what can be said about the phenomenon *given our current knowledge*. Positions (1.) and (3.), then, are stronger than their counterparts, but are scientifically and epistemologically less sound because they only exist as logical possibilities and not as actual theories. In other words, it is never possible to speak about the world itself without making reference to the theories we use to describe it, thus an absolute position fails when subjected to sceptical remarks. Secondly, it is clear that for an event A to be epistemologically reducible to B the ontological reducibility of A to B is a necessary condition. However, the reverse implication does not hold: the ontological reducibility of a phenomenon to another does not imply that no further explanation is needed.

When these categories are applied to the problem of neural determinism a multiplicity of possible interpretations of the problem arise, which I do not delve into. In a simplistic description of the theoretical background on the subject, the two main possibilities are reductionism and antireductionism. Reductionism is the view according to which mental processes are entirely reducible to neurological processes, from both an ontological and an epistemological perspective. Antireductionism is that according to which mental processes are not entirely reducible to neurological ones, either from an ontological point of view or an epistemological one. A reductionist, then, might believe that, since we are made of atoms, it is ontologically and explicatively sufficient to explain how these atoms interact with each other and the others surrounding them to make sense of human thought and behaviour. An anti-reductionist might either believe that there is something other than our atoms which forms our mind or that giving a complete description of our atoms and their interactions does not provide a satisfactory explanation of our behaviour, its causes and its mechanisms. The latter case is quite popular as it can be linked to the *hard problem of consciousness*, i.e. with the problem of why there is such a thing as *first-person experience* in our consciousness. This problem is that of qualitative experience: it is possible to describe everything that is happening in a subject's brain, but this would not tell us *how* the subject is experiencing what she is experi-

19 - It is extremely important that here I refer to our best quantitative models of a phenomenon, and not of the phenomenon itself.

encing. Notably, following T. Nagel's[20] example, it is possible to describe everything that happens in a bat and still have no idea of *what it is like to be a bat*.

Finally, another important theme to point out is that the notion of consciousness does not theoretically exhaust what can be said about the relationship between thought and neural determinism. While it is clear that consciousness is a brain-bound and neurally determined process, there are other conceptual notions that have to be employed in describing thought, such as that of cognition, which do not have the same characteristics. The notion of consciousness is usually employed because it is intrinsically linked with that of personal identity and phenomenological experience. However, it is also useful to use the concept of cognition, because it allows us to dive deeper into the relationship between neural determinism and biological determinism as a whole. Cognition has been analysed in terms of internalism, embodiment and extension, thus moving away from the pure contribution of the brain in our mental processes and into the influence of external factors. Following a variety of approaches on cognition[21], in order to consider and understand our thought processes it is not sufficient to focus on the brain and its functioning, but it is necessary to widen the scope of the inquiry to the consideration of bodily processes and the out-sourcing of bodily functions that can be obtained with the construction of ecological niches and the cultural scaffoldings.

*Free will or personal agency?*

A notable feature of Dostoevsky's understanding of the problem of neural determinism is that it is inherently linked with its ethical component: the repercussions of determinism on moral responsibility. Intuitively, a reasoner is morally responsible for her actions if she has performed them freely and willingly.

There are at least three necessary conditions[22] to talk about an action as performed freely:

1. The ability to do otherwise: in order for a person to be free she has to concretely have the opportunity to decide differently from what she does.
2. Agency: for a person to act according to her own free will she has to be in control of her actions. Her actions need to be up to her.
3. Responsiveness to reasons: the action, to be regarded as free, cannot be automatic and unconscious, but has to be put under rational scrutiny.

On top of having chosen freely, to be held accountable for her actions a reasoner has to have understood what she was choosing and has to have deliberately decided that she wanted to perform it.

An overview of this kind becomes extremely problematic within the theoretical framework of biological determinism because determinism undermines the ability of a reasoner to choose freely and willingly. Apparently, this would lead us to believe that Dostoevsky feels there is an incompatibility between neurological determinism and free will, as determinism seems to undercut the first criterion of free choice (i.e. the ability to do otherwise). However, it has to be noted that there are at least two objections to this conclusion. First of all, empirical evidence has been gathered that supports the idea that there is a very strong cultural bias in the tendency to accept the first criterion of free choice: often enough an incompatibility between moral responsibility and the inexistence of alternative courses of action is not felt. In particular, populations of Asian descent, for instance, statistically do not seem to feel a strong contrast between complete determinism and individual responsibility, which, on the contrary, is strongly felt in Western culture[23]. Secondly, while determinism is often regarded as incompatible with free will, there are many researchers who support versions of *compatibilism*[24]. According to compatibilism, the

20 - Nagel, Thomas. "What is it like to be a bat?" Philosophical Review, 83, 1974, pp. 435-50.

21 - See, for example: Clark, Andy and D. Chalmers, Daniel. "The extended mind". Analysis, 58, 1998, pp. 7–19 ; Clark, Andy. "Supersizing the Mind". Oxford University Press, 2010 ; Menary, Richard. "The Extended Mind". The MIT Press, 2010 ; Menary, Richard. "Growing minds". Habits, Cambridge University press, 2020, pp. 297-319.

22 - Lavazza, Andrea. "Free Will and Neuroscience: From Explaining Freedom Away to New Ways of Operationalizing and Measuring It." *Frontiers in human neuroscience* vol. 10 262. 1 Jun. 2016, doi:10.3389/fnhum.2016.00262.

23 - Hannikainen, Ivar R et al. "For Whom Does Determinism Undermine Moral Responsibility? Surveying the Conditions for Free Will Across Cultures." *Frontiers in psychology* vol. 10 2428, 2019, doi:10.3389/fpsyg.2019.02428.

24 - McKenna, Michael and D. Justin Coates, "Compatibilism", The Stanford Encyclopedia of Philosophy (Fall 2021 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/fall2021/en-

criterion of alternative possibilities is too strong since free will can be compatible with determinism (creating a sort of *soft* determinism). These objections help differentiate more subtly what constitutes the moral and philosophical problem in Dostoevsky's eyes: not free will, but rather individual agency.

Specifically, Dostoevsky does not speak of determinism as involving a loss of freedom, but rather, paradoxically, the contrary. It is true that the author constantly juxtaposes the image of neurological determinism with that of prisoners, but at the same time Dmitri talks about the new man, the man born out of science and atheism, as free and absolutely responsible for all of his actions ("*Without God and the future life? It means everything is permitted now, one can do anything?[...] Everything is permitted to the intelligent man*"[25]). The emphasis, then, is not on how determinism corrupts the possibility of freedom itself. What is invalidated by determinism is rather the possibility of being in control of our own thoughts. The problem is not that men cannot choose to do something, but rather that their thoughts themselves are not up to them and are completely shaped by biology and environment. This leads Dmitri to his cry and sorrow: we think not because of the agency we find in our personal identity, but because of the neuronal happenings within our brains. What is at stake, then, is not just the possibility of moral responsibility and free choice, but the existence of personal identity itself as traditionally conceived. The dilemma of free choice follows, of course, from this lack of individual agency.

*Dostoevsky's paradigm: emotion as a way of faith*

The importance of the theme at hand seems to warrant a thorough analysis and argumentation, which is what we would expect from Dostoevsky. The author, however, as shown in the cited passage, does nothing of the kind, and simply relies on an appeal to emotion (Dmitri's sorrow towards the abandoned God) in order to contrast the consequences of determinism. Why exactly is that?

Dostoevsky could have rationally and philosophically refuted the argument for biological determinism in the following ways:

- By invoking the strategy of the so-called *"God of the holes"*[26]: this strategy relies on the fact that, since there is always something that is not explained by the current scientific knowledge, God must be present in that unexplained data. This strategy, however, is weak because it is hostage to contingencies: with the further encroachment of science the space left for God would become increasingly small.
- By simply adopting a sceptical approach on the scientific theories of his time. Even though there was a growing body of scientific research on cognitive and neurological phenomena, these processes were still almost completely unknown. Golgi had just invented his black reaction method to show neurons and Cajal had not yet developed his neuronal theory. The knowledge that was available on epilepsy itself was rudimentary, and it often bordered on superstition[27]. Nevertheless, it is clear that it would have not been a rational and scientific approach to simply dismiss the research of the time in spite of its empirical success.
- By accepting one of the possible forms of the previously mentioned anti-reductionism that would have been available to him (i.e. he could not have thought of the role of qualia per se, but it would have been possible for him to posit the existence of extra-physical entities or the ineliminable qualitative difference between physical and mental processes). This strategy is quite interesting because, in some forms, it is still accepted. However, Dostoevsky would not have had the theoretical tools needed to obtain a strong anti-reductionism and his argument would have begged the question.

Dostoevsky, however, chooses not to pursue these lines of argumentation, which, as shown, are extremely weak, and favours an approach where biological/neurological determinism is readily accepted, but met

tries/compatibilism/>.

25 - Dostoevsky, Fyodor. "The Brothers Karamazov". Translation by Richard Pevear and Larissa Volokhonsky, 1992, pp. 499-500.

26 - Ratzsch, Del and Jeffrey Koperski, "Teleological Arguments for God's Existence", The Stanford Encyclopedia of Philosophy (Spring 2022 Edition), Edward N. Zalta (ed.), forthcoming URL = <https://plato.stanford.edu/archives/spr2022/entries/teleological-arguments/>.

27 - Iniesta, Ivan. "Epilepsy in Dostoevsky", Editor(s): Anne Stiles, Stanley Finger, François Boller, Progress in Brain Research, Elsevier, Volume 205, 2013, Pp. 277-293.

with emotional sorrow. This constitutes a novel and unexpected paradigm: Dmitri's faith does not reject the findings of science. He accepts them and believes nonetheless.

Is this strategy philosophically rational and viable? Most abstract models[28] on rationality dismiss the role of emotions, stating that they are intrinsically irrational and, thus, should not be part of good decision making processes. Nevertheless, starting from the early 2000s[29] and more and more in the last few years, recent literature on decision making has given emotion a second chance on the basis of its fundamental evolutionary and ecological role. The thesis of the bounded approach on rationality[30] is that what can be considered rational is not what models abstractly developed *a posteriori* tell us on how we ought to reason, but rather the conditions that are ecologically relevant in the context within which reasoning norms have emerged. Thus, emotion is tightly bound with reasoning and cannot be expunged from it, in the same way social and cultural factors cannot be completely expunged from the consideration of scientific theories. These emotions, however, cannot be simply accepted in virtue of their existence, but need to be first analysed and understood.

Given the most recent developments in the research of the relationships between reason and emotion, thus, Dostoevsky's paradigm seems to be philosophically viable, provided that these emotions be properly understood and employed.

## Conclusion

Dostoevsky employs his personal experience and his knowledge of brain processes and diseases in giving a clear picture of his conception of the MB problem and the consequences of biological determinism on human action. Despite the fact that some conceptual differences - such as the layering of the problem of biological determinism, the role of embodiment in determinism, or the various possible anti-reductionist strategies - were not available to him, he developed a picture of the notions of neurological determinism and human agency which was clear and incredibly advanced for his time. Particularly interesting is the development of a paradigm of faith which still holds, renewed by recent research, and is intellectually sound in that it does not refute the data from hard sciences while still trying to bridge the gap with soft sciences.

28 - Kahneman, Daniel. "Thinking, Fast and Slow". New York: Farrar, Straus and Giroux, 2011.

29 - Hanoch, Y. (2002). "Neither an angel nor an ant": Emotion as an aid to bounded rationality. *Journal of Economic Psychology, 23*(1), 1–25. https://doi.org/10.1016/S0167-4870(01)00065-4.

30 - Viale, Riccardo. "Routledge Handbook of Bounded Rationality". 2021.

## References


- Bernard, Claude. "Leçons sur les Phénomènes de la Vie", in Conti, F. "Claude Bernard's Des Fonctions du Cerveau: an ante litteram manifesto of the neurosciences?", in Nature Reviews of Neuroscience, 2002.
- Brigandt, Ingo and Alan Love, "Reductionism in Biology", The Stanford Encyclopedia of Philosophy (Spring 2017 Edition), Edward N. Zalta (ed.), <https://plato.stanford.edu/archives/spr2017/entries/reduction-biology/>.
- Clark, Andy. "Supersizing the Mind". Oxford University Press, 2010.
- Clark, Andy and Chalmers, Daniel. "The extended mind". Analysis, 58, 1998, pp. 7–19.
- Cohen, Sharon. "'Balaam's Ass': Smerdyakov as a Paradoxical Redeemer in Dostoevsky's The Brothers Karamazov." Christianity & Literature, vol. 64, no. 1, Dec. 2014, pp. 43–64, doi:10.1177/0148333114552772.
- Dostoevsky, Fyodor. "The Brothers Karamazov". Translation by Richard Pevear and Larissa Volokhonsky, 1992.
- Feyerabend, Paul. "Dialogo sul Metodo". Translation by R. Corvi, Edizioni Laterza, 2007.
- Hannikainen, Ivar R et al. "For Whom Does Determinism Undermine Moral Responsibility? Surveying the Conditions for Free Will Across Cultures." *Frontiers in psychology* vol. 10 2428, 2019, doi:10.3389/fpsyg.2019.02428.
- Hanoch, Y. (2002). "Neither an angel nor an ant": Emotion as an aid to bounded rationality. *Journal of Economic Psychology, 23*(1), 1–25. https://doi.org/10.1016/S0167-4870(01)00065-4.
- Hoefer, Carl, "Causal Determinism", The Stanford Encyclopedia of Philosophy (Spring 2016 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/spr2016/entries/determinism-causal/>.
- Iniesta, Ivan. "Epilepsy in Dostoevsky", Editor(s): Anne Stiles, Stanley Finger, François Boller, Progress in Brain Research, Elsevier, Volume 205, 2013, Pp. 277-293.
- Kahneman, Daniel. "Thinking, Fast and Slow". New York: Farrar, Straus and Giroux, 2011.
- Lavazza, Andrea. "Free Will and Neuroscience: From Explaining Freedom Away to New Ways of Operationalizing

and Measuring It." *Frontiers in human neuroscience* vol. 10 262. 1 Jun. 2016, doi:10.3389/fnhum.2016.00262.
- McKenna, Michael and D. Justin Coates, "Compatibilism", The Stanford Encyclopedia of Philosophy (Fall 2021 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/fall2021/entries/compatibilism/>.
- Maze, J. R. "Dostoyevsky: Epilepsy, Mysticism, and Homosexuality." American Imago, vol. 38, no. 2, The Johns Hopkins University Press, 1981, pp. 155–83, http://www.jstor.org/stable/26303748.
- Menary, Richard. "The Extended Mind". The MIT Press, 2010.
- Menary, Richard. "Growing minds". Habits, Cambridge University press, 2020, pp. 297-319.
- Mourad, Ronney. "Choosing to Believe." International Journal for Philosophy of Religion, vol. 63, no. 1/3, Springer, 2008, pp. 55–69, http://www.jstor.org/stable/27646256.
- Nagel, Thomas. "What is it like to be a bat?". Philosophical Review, 83, 1974, pp. 435-50.
- Ratzsch, Del and Jeffrey Koperski, "Teleological Arguments for God's Existence", The Stanford Encyclopedia of Philosophy (Spring 2022 Edition), Edward N. Zalta (ed.), forthcoming URL = <https://plato.stanford.edu/archives/spr2022/entries/teleological-arguments/>.
- Sapolsky, Robert. "Behave: The Biology of Humans At Our Best and Worst". New York, New York: Penguin Press, 2017. Print.
- Viale, Riccardo. "Routledge Handbook of Bounded Rationality". 2021.
- Wittgenstein, Ludwig. "Philosophical Investigations". Oxford, UK: Blackwell, 1997. Print.