(Brain) Death, (Brain) Life and the Value of Life

Christian Erk*

University of St. Gallen, Switzerland

Abstract

If we accept the “Dead-Donor-Rule (DDR)” as a true and valid premise, we are immediately confronted with the arduous task of answering one of the most controversial questions of contemporary bioethics: How can we diagnose an explantee’s death with certainty? Taking this problem as its starting point this article takes a critical look at and suggests expanding the so-called “definition-criteria-tests (DCT)” model that is usually used to answer the question just posed. This suggestion derives from the insight that – as will be argued – every definition of death is necessarily based on a definition of life that logically precedes it. Applying the refined DCT model to the whole brain definition, the article then sets out to reveal the strange notion of life implied by brain-based definitions of death and by doing so questions their tenability. Accepting that brain-based approaches are not suitable to define death would allow for a concentration on the problem that really is at stake in the debate about the (im)permissibility of the transplantation of vital organic material, namely our understanding of the phenomenon we call life and its value.

Keywords: (brain) death, value, value of life, (brain) life

INTRODUCTION

When moral philosophers argue for or against the permissibility of the transplantation of vital organic material, i.e. any organic material that is indispensable to (the explantee’s) life, their argument is – at least implicitly – based on one of the two following mutually exclusive premises: They either accept or repudiate the so-called “Dead-Donor-Rule (DDR)”. According to this rule, any explantation of organic material from a living human person $P_1$ (explantee) destined for the subsequent implantation into another living human person $P_2$ (implantee), that results in and causes the death of the explantee, is morally impermissible. The DDR, in other words, states that the explantation of any vital organic material for transplantation from an explantee is morally impermissible if the explantee has not been declared dead before the beginning of the explantation.

Although there are some writers (e.g. [1]) who (in most cases based on utilitarian reasoning) suggest abandoning the DDR, let us – for the purposes of this article – assume that the DDR is indeed a true and valid premise. If we do so, we are immediately confronted with the
arduous task of making up our minds about one of the most controversial questions of contemporary bioethics: How can we know and diagnose with certainty that an explantee (or generally speaking: a human person) is dead?

Taking this problem as its starting point this article shall take a critical look at and suggest expanding the so-called “definition-criteria-tests (DCT)” model that is usually used to answer the question just posed. This suggestion derives from the insight that – as will be shown – every definition of death is necessarily based on a definition of life that logically precedes it. Applying the refined DCT model to the whole brain definition, the article then sets out to reveal the strange notion of life implied by brain-based definitions of death and by doing so questions their tenability. Accepting that brain-based approaches are not suitable to define death would allow for a concentration on the problem that really is at stake in the debate about the (im)permissibility of the transplantation of vital organic material, namely our understanding of the phenomenon we call life and its value.

THE DCT MODEL FOR THE FORMULATION OF DEATH

In order to address the problem of the diagnosis of death, contemporary ethical debate usually draws on the so-called “definition-criteria-tests (DCT)” model. According to this model, “every formulation of death has three components: a concept or definition of what it means to die; operational criteria for determining that death has occurred; specific medical tests showing whether or not the criteria have been fulfilled.” [2, p. 200] According to this model, diagnosing death means applying the criteria and tests that operationalise a given definition of death: While the tests operationalise the criteria, the criteria are a mere operationalisation of some definition of death.

Since the operational criteria and specific medical tests only serve to confirm whether some definition of death has been fulfilled, it should be beyond dispute that the definition of death enjoys logical primacy within the DCT model: One cannot start thinking about the criteria and tests of death except on the basis of a previously given definition of death. Therefore, even the best criteria and tests of death are useless if the definition of death they derive from is unsound. The opposite is, however, not true: The quality and acceptability of a definition of death does neither depend on the quality of its corresponding criteria and tests nor on their existence.

This means that when it comes to the diagnosis of death ensuring the quality of one’s definition of death is of the utmost importance. So, what does the quality of a definition of death depend on? On a general level, the answer to this question should be rather straightforward: The quality and acceptability of a definition of death only depends on how well it answers the question “What does it mean to die?” and, therefore, the question “What is death?”. And it is exactly when thinking about the latter question that we should realise that the three-level DCT model is insufficient and should be complemented.

DEATH AS THE ABSENCE OF LIFE

What is it that the DCT model is lacking? The DCT model’s problem is that it requires us to define something we cannot define in a direct or positive fashion. How so? One of the first things we (should) notice when addressing the question “What is death?” is that we are somewhat barred from gaining any substantial insights into the occurrence we call death. Finding a definition of death is somewhat foredoomed by the fact that we cannot make any direct or positive substantial statement about death. We know that death marks the end of a process we
call dying and that it is certain. But, soon after these rather superficial insights our inquisitive efforts hit a (natural) roadblock.

Death is not something we can observe and define in the same fashion as we can observe and define e.g. a material object or even such a thing as an emotion. For, there is only one way to experience death: by actually dying. The problem is that those of us who do experience and observe death in this way cannot tell us about it. As an object of investigation death is drenched in a kind of darkness that is impenetrable to the eye of the living; it eludes our attempts to get a direct or positive (scientific or philosophical) grasp of it.

This does, however, not mean that there is nothing substantial we can say about death. For though we cannot make any direct or positive statement about the essence of death (and the essence is what is supposed to be captured by a definition), there is something substantially to be said about death in an indirect or negative fashion: Death is the absence of something (X) that has been present before and that is lacking after death has occurred. Regardless of its actual content, the nature of death as an absence or negation is something that is at least implicitly and unconsciously recognised by all definitions of death – be they brain, cardiopulmonary or otherwise based. So, the only way we can actually answer the question “What is death?” is by saying: “Death is the absence of X.”

But: What is this X that is supposed to be present before and absent in death? The most straightforward approach to answer this question is to look for death’s contradictory opposite, i.e. that which is incommensurate with death. Although not part of the world of the living, death has a contradictory correlate in the world of the living, namely life itself. Life and death are mutually exclusive categories: Something that lives cannot be dead; and something that is dead, cannot be alive. Regardless of what else we take life to be, being dead simply means having lost one’s life. Ultimately, the first and most fundamental answer to the question “What is death?” and thus the only general definition of death we can give is the following: Death is the absence, i.e. the completed and irreversible loss, of life.

Some might object that the same that is true for life (i.e. that it is present before and absent after death) also applies to e.g. brain activity or cardiopulmonary activity or the capacity to smile. But, although all of this is indeed present before and absent after death, death can be defined as the absence of these (or some other) activities or capacities only if and insofar as they are taken to be constitutive of life.

We should not commit the logical impossibility to look at life through the lens of death; what we should, rather, do is to look at death through the lense of life. While our understanding of death necessarily hinges upon a previous definition of life, what life is has to be determined independently from our understanding of death. When it comes to defining death there are two lines of argument which are both based on the same second premise (P2), namely the general definition of death as the absence of life, but differ with respect to the first premise (P1) and the conclusion (C). These lines of argument read as follows:

<table>
<thead>
<tr>
<th>Line of Argument A</th>
<th>Line of Argument B</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1: Life is Y.</td>
<td>P1: Death is the absence, i.e. the complete and irreversible loss, of X.</td>
</tr>
<tr>
<td>P2: Death is the absence, i.e. the complete and irreversible loss, of life.</td>
<td>P2: Death is the absence, i.e. the complete and irreversible loss, of life.</td>
</tr>
<tr>
<td>C: Death is the absence, i.e. the complete and irreversible loss, of Y.</td>
<td>C: Life is X.</td>
</tr>
</tbody>
</table>

According to line of argument B, the definition of life derives from an underyling definition of death. Given the fact that death’s contradictory correlate is part of the real world of the
living while death is not, this approach seems to be rather odd and error-prone. For, the definition of death given in P1 has to be put to the test by verifying the resulting definition of life by the yardstick of our intuition and common experience. If the resulting definition of life is contrary to our intuition and common experience, then the definition of death given in P1 has to be refined – until one finally ends up with a definition of life that fits.

Instead of putting the cart before the horse, i.e. instead of trying to define death before having defined life, it would make much more sense to start with a definition of the phenomenon that is not an absence but a presence, namely life.

AN EXTENDED MODEL FOR THE FORMULATION OF DEATH

Embracing the insight that the most fundamental definition of death is that is the absence of life has, however, grave implications. For, it reveals that the DCT model asks us to take the second step without having taken the first. If death first and foremost is the absence of life, this means that every definition of death must be preceded by a definition of life because every definition of death just is the negation of a definition of life. As the definition of death enjoys logical priority in comparison to the criteria and tests of death, so does the definition of life in comparison to the definition of death. Any complete formulation of death must start with and emanate from a definition of life, i.e. an answer to the question “What is life?”.

Starting with a definition of death means starting in the middle of the argument without explicating one’s premises.

In consequence, the three-level DCT model for the formulation of death should be complemented by an additional level, namely the definition of life. The model that should guide our thinking about the diagnosis of death should, therefore, be comprised of the following four levels respectively deliberation steps:

1. Definition of Life (DL)
2. Definition of Death (DD)
3. Criterion/ Criteria of Death (C)
4. Test(s) of Death (T)

Adding a fundamental level and by doing so changing the DCT model into this “definition of life-definition of death-criteria-tests (DLDDCT)” model has, however, certain consequences for the way we think about death. For, by starting our endeavours to diagnose death from a definition of life we should be constantly reminded that what we are really talking about is not (the diagnosis of) death but (the diagnosis of) the absence of life. In reality, what we have called criteria of death actually are criteria of the absence of life and what we have called tests of death actually are tests of the absence of life.

However, establishing criteria of the absence of life is not as straightforward as it might seem at first sight. For, if one wants to know how to determine the absence of something one first has to know how to determine its presence. This means that in order to be able to establish a criterion or set of criteria (sufficient and necessary) of the absence of life one has to first understand how to determine life’s presence. When it comes to the criteria of the absence of life one has to, therefore, answer the following two questions in the order presented:

• By means of which suitable criterion or set of criteria (sufficient and necessary) can the presence of life be determined (certain and uncertain signs of (the presence of) life)?
• By means of which suitable criterion or set of criteria (sufficient and necessary) can the completed as well as irreversible loss and, therefore, absence of life be determined (certain and uncertain signs of the absence of life, i.e. death)?
Only after these questions have been answered in a satisfactory fashion, one can set out to compile a list of specific medical tests that are suitable to reliably show whether or not the criteria have been fulfilled.

Summing up what has been said so far allows us to improve on and expand the initial formulation of the DL_DoCT model as follows:

1. **Definition of Life (D_L)**
   
   *Guiding question:* What is life?

2. **Definition of Death (D_D)**
   
   The *guiding question* of this level of the DL_DoCT model ("What is death?") is already answered as death is the absence, i.e. the completed and irreversible loss, of life. So, death is the absence of whatever is defined as life.

3. **Criteria of the Presence and Absence of Life (C)**
   
   *Guiding questions:*
   
   - By means of which suitable criterion or set of criteria (sufficient and necessary) can the presence of life be determined (certain and uncertain signs of (the presence of) life)?
   - By means of which suitable criterion or set of criteria (sufficient and necessary) can the completed as well as irreversible loss and, therefore, absence of life be determined (certain and uncertain signs of the absence of life, i.e. death)?

4. **Test(s) of the Presence and Absence of Life (T)**
   
   *Guiding question:* By means of which validated and agreed-upon confirmatory medical test procedure(s) and parameter(s) can be reliably shown whether the criteria of the presence or absence of death are fulfilled?

   Only after these four steps have been taken and only after all these questions have been explicitly addressed it is possible and reasonable to diagnose the death of a human being. Unless the criteria and tests used are a logically consistent operationalisation of an explicitly given definition of life, every diagnosis of death is arbitrary. This also means that, regardless of its actual content, every definition of death is only as meaningful and tenable as the definition of life it implicitly or explicitly negates.

   So, what the debate on the (im)permissibility of the explantation of vital organic material should really be about is not the question “When has death occurred?” but rather the question “When has life ended?” For, death can only be fully understood if looked at through the lense of life.

---

1 Unfortunately, the contemporary debate on the (im)permissibility of the explantation of vital organic material does not follow this advice but tries to argue about the absence of something without having previously defined what this something actually is. Virtually none of the articles I have come across so far explicates the definition of life that the definition, criteria and tests of death it advances are necessarily based upon. It has to be admitted that there are articles that advance or at least discuss the concept of “brain life” as opposed to “brain death” (cf. [3, p. 199; 4, p. 475ff; 5, p. 50; 6, p. 14; 7; 8, p. 122; 9, p. 439; 10, p. 33f; 11, p. 76f; 12, p. 44ff]). However, none of these articles actually offers a definition of life that goes beyond seeing life as the mere absence of death. Instead, what these articles do is applying their respective notion of death to the problem of the beginning of life. It is time and a matter of intellectual honesty to remedy this unfortunate situation and start focussing on the real issue at stake, namely life.
Rephrasing the problem in this fashion accentuates the fact that whatever ends must have also had a beginning. The life that ends in death must have come into being beforehand; our lives have not only one but two boundaries. And these boundaries are logically interdependent: The way we draw the boundary that marks the end of life predetermines the way we have to draw the boundary that marks its beginning; and vice versa. Therefore, our definition of life is not only decisive for our answer to the question “When has life ended?” but also for our answer to the question “When does life begin?”.

In other words: If we want to be logically consistent, the criteria and tests we use to determine the point in time when life ends must also be used to determine the point in time when life begins.

What happens if we apply this requirement to the definition of death that seems to dominate contemporary philosophical debates and transplantation policies, namely the whole brain definition of death? According to the D1D2CT model for the formulation and diagnosis of death, every definition of death merely is the negative formulation of a definition of life that it is necessarily based on. So, what is the definition of life that the whole brain definition of death is the negation of? According to this definition, death is the irreversible cessation of all brain activity, i.e. the irreversible loss of the functioning of the entire brain, including the brain stem. If death is defined as the absence of life and if the whole brain definition defines death as the cessation of all brain activity, this allows us to conclude that the definition of life that underlies the whole brain definition of death is that life consists in the presence of brain activity. Subscribing to the whole brain definition of death means subscribing to the position that human beings are alive not only as long but also as soon as they exhibit brain activity.

However, subscribing to the whole brain definition of death means subscribing not only to the position that human beings but that all beings are alive not only as long but also as soon as they exhibit brain activity. Accordingly, a proponent of the whole brain definition of death has to not only defend the position that brain-dead human beings are dead. What’s more, he has to – for the sake of consistency – also support the position that all prokaryotes (e.g. bacteria), all plants, all fungi, all animals (i.e. also all human beings) in the early stages of their development and some simpler animals over the course of their entire existence (e.g. sponges) are not alive; for, none of these beings does and can exhibit brain activity – because none of these has a brain. It would be interesting to see whether any proponent of the whole brain definition of death is willing to actually support this highly counterintuitive position that logic forces him to support.

However, let us, for the sake of convenience, leave aside all the other beings just listed and have a closer look at human beings in the early stages of their development. According to the whole brain definition of death, at what point in time during their development do human beings actually become alive? Answering this question amounts to finding an answer to the question of when human beings start to exhibit brain activity. The earliest point in time it is not impossible for them to do so is once at least some basic brain-like structures have developed and start to function. As embryology tells us, the development of the structural pre-

---

2 The same thought experiment can, of course, also be conducted for the partial brain definitions of death (higher brain death; brain stem death; etc.) as well as the cardio-pulmonary definition of death. The definition of life that forms the basis of e.g. the cardio-pulmonary definition of death (which equates death with the complete and irreversible loss of all cardiopulmonary function, i.e. function in the heart and lungs) is that life is the presence of heart and lung activity.

3 A proponent of this “brain-life theory” can be found in [3, p. 200]: “Whenever a functioning human brain is present, a human being is alive.”
The first cursors of a human being’s brain starts with a process called “primary neurulation”. This process is normally initiated 19 days after conception (Carnegie stage 7) with the formation of the so-called “neural plate”, i.e. the key developmental structure that serves as the basis for the entire nervous system. It is, however, questionable whether the neural plate can be said to exhibit brain activity that can be detected and measured by means of an electroencephalography (EEG). Such electrocerebral activity can usually be measured around week 8 (gestational age) and, therefore, around week 6 (fertilisation age) of an embryo’s development (Carnegie stage 19). If it is only at 8 weeks post menstruationem respectively 6 weeks post conceptionem that the brain of the embryo first generates electrical waves (brain waves) discernible by an EEG, we are led to conclude that a proponent of the whole brain definition of death has to accept the position that human beings are definitely not alive during the first 6 weeks of their existence.

If this were true this would mean that a pregnant woman carries and nourishes a dead being in her womb for the first two months of the pregnancy, that human embryonic stem cell researchers hope to engineer tissue from dead cells, that the human beings who come into existence by means of in-vitro fertilisation are “conceived” dead, that the embryos transferred into the mother’s uterus after an in-vitro fertilisation are dead embryos and that pre-implantation (genetic) diagnosis is performed on dead human beings. What’s more, if all this were true, this would mean that human embryos cannot be killed (at least not during the first 6 weeks of their existence) because it is impossible to kill a being that is not alive. In other words: If we were to accept the whole brain definition of death, it would be superfluous to argue about the (im)permissibility of embryocide, because there would be no such thing as embryocide.

**THE REAL CHALLENGE POSED BY BRAIN-BASED DEFINITIONS OF DEATH**

As biting this bullet and accepting these (to put it mildly) highly counterintuitive notions would not be a reasonable thing to do, proponents of the whole brain definition of death often argue that their definition of death is not a definition of actual and complete death but is only meant to signify (a) the loss of that aspect of life that is morally relevant or (b) the loss of that capacity the possession of which is the taken to be the precondition for moral status and personhood. According to this line of reasoning, brain death is not the death of the human being as a living human being but as a being with moral status. In other words: A brain dead human being is (at least partially) alive but lacking moral status or personhood. However, this line of argument cannot be advanced without conceding that a brain-dead human being is not entirely dead but both partially dead and alive. And doing so means admitting that the whole

---

4 This is also confirmed by e.g. [13, p. 44] who states that brain waves first appear “at about 8 weeks g.a.” (also cf. [9, p. 436; 14]). For the sake of completeness it has to be annotated that the brain activity detectable in Carnegie stage 19 does not yet resemble that of a newborn let alone a grown up human being. The brain wave pattern of the unborn start to resemble the regular brain wave pattern of a newborn some 27 to 28 weeks post conceptionem. While this point in time might be relevant for proponents of higher brain definitions of death, it is, however, irrelevant for the whole brain definition of death that defines death as electrocerebral silence, i.e. the cessation of all electrocerebral activity (as established by an isolelectric EEG (also: zero line EEG; flat EEG)) and not just activity in form of regular brain waves.

5 This is usually done on the basis of a division of life into biological, sensitive and rational life or – as in [15] – into ζωή (read: zōē) and βίος (read: bios)).

6 Although the exact terminology with respect to this capacity varies (capacity for thought, self-awareness, capacity to have interests, …), the most general label for this capacity, i.e. the label under which virtually all terms can be subsumed, is “rationality”.

7 cf. McMahan who speaks of the death of the person and the death of the organism (cf. [9, p. 439; 8, p. 122]).
brain definition of death is not a definition of actual and complete death but only explains one aspect of it\(^8\). Admitting this would be of the utmost importance because it would allow us to tackle the real ethical issues that are at stake with regard to the ex- and transplantation of vital organic material. At its core, the debate about the (im)permissibility of the transplantation of vital organic material and the debate about (un)tenability of the brain death definitions of death that is contained in it are not about the moment of actual and complete death; rather, although well disguised, they is about (human) life and its value as well as the moral status or personhood of living (human) beings. It is time to realise this and rise up to the real challenges presented to us by the emergence of the brain-based definitions of death, namely: (1) to explain what (human) life is\(^9\), (2) to show if, why, under what conditions, from what moment on and how long (human) life possesses value, (3) to reason how this value is to be respected and (4) to explain whether a human being’s life can weighed against the life of another human being. But with the definition of death, one is not free to tackle these questions in random order. The second, third and fourth question can only be addressed once the first question has been answered. Answering question (1) is of such importance that explicating one’s definition of life should be the obligatory entry ticket to participating in any debate about issues of life and death – regardless of whether they concern the beginning or the end of life.

REFERENCES


\(^{\text{8}}\) This conclusion is also perfectly true for all partial brain definitions of death.
\(^{\text{9}}\) In an attempt to answer this question I have presented a definition on life elsewhere: [16, p. 162ff].