Abortion: What We Ought to Believe.

An Ontological and Normative Analysis.

José Ezequiel Páez Conesa

TESI DOCTORAL UPF / 2013

DIRECTORS DE LA TESI

Dra. Marisa Iglesias i Vila

Dr. José Juan Moreso i Mateos

DEPARTAMENT DE DRET

For my grandmother Isabel,
since she always strives to do what she deems best.

'Here the imagination of man has been very busy. Vain speculations have existed as to the moral nature; yea, even as to the great question of the moral accountability of unborn children. The medical philosopher, or rather metaphysician (for the boundaries of true philosophy have been transgressed), and also the professed theologian, have given reins to their imagination […].'

Hugh L. Hodge, M.D., 1839[1]

[1] Hodge (1869, p. 20).

Acknowledgments

When I started working on this dissertation I had almost no knowledge of moral philosophy or metaphysics. Yet as the subtitle of this work indicates, these are the two fundamental disciplines from which the problem of abortion ought to be analysed. Thus, it is not out of mere courtesy, but because it is strictly true, that I acknowledge that the product of my endeavours would have been much more deficient without the guidance or contributions of many people. I will try to do them justice here. If there is anyone absent who should be present, please do not construe it as a sign of ingratitude, but as the failings of a forgetful mind. Obviously, these people are entirely guiltless of any mistakes or infelicities in the following text. Those are exclusively my own fault.

Let me start with my colleagues. I had the tremendous fortune of becoming a doctoral candidate in one of the finest, if not the finest, research groups in practical philosophy in the country. My gratitude is first due, then, to the members of the Legal Philosophy Area at Pompeu Fabra University. They provided me with a superb work environment for philosophical exploration and intellectual growth.

Out of all of them two deserve to be especially thanked. These are my supervisors, Marisa Iglesias and José Juan Moreso: they were kind enough to support my entrance into the group and, later, to trust my capacity to pursue this line of research. For the last five years José Juan, as the university rector, struggled to find a spot in his tight schedule to discuss my progress. For her part, Marisa has been the best supervisor I could have hoped for: generous with her time, a sharp discussant, and a judicious administrator of both encouragement and criticism. Her comments on each part of the dissertation and the document as a whole were always extremely helpful. It was soothing to know that I could always come to her in times of need.

Certainly, though, theirs was not the only assistance which I received from the colleagues at the university, and for which I am thankful. Serena Olsaretti

provided me with detailed criticism of a draft of Chapter I, which I could consequently greatly improve. Paula Casal, José Luis Martí, Hugo Seleme and Andrew Williams made very valuable suggestions over all these years, both before and after the actual writing process began. It is amazing what you can learn at lunch with the right people if only you pay attention.

Though my topic was far removed from their area of expertise, there are many others from whose conversation I have greatly benefited. I have in mind, of course, such wonderful persons as Jorge Malem, Leticia Morales, José Luis Pérez Triviño, Neus Torbisco and Josep Maria Vilajosana. I want also to thank the latter for his course on argumentation: I have always tried to bear his teachings in mind, even though I am sure I have not completely succeeded at applying them. Two other professors deserve to be mentioned here for their courses. The first of these is Ricardo Caracciolo: I shall never forget his fascinating seminars on the work of Joseph Raz. The second, but not the least of them, is Ernesto Garzón Valdés. I have learned a great deal in his seminars about the history of philosophy, and his academic life has always appeared to me to be a source of inspiration.

I must also thank the people from the Uehiro Centre for Practical Ethics at the University of Oxford, where I stayed from April to June in 2011. There I had the chance to observe philosophical work in progress which met the highest expectable standards. Julian Savulescu, the Centre's director, deserves especial gratitude, for he reserved part of his time to discuss my project with me and to encourage me to pursue it.

I now turn fondly to those members of the Area which constituted the Law and Philosophy Meetings (LPM) discussion group, to which I belong, and which devoted many of its earliest sessions to the consideration of several chapters of this dissertation. Since the idea of the LPM must be credited to Sebastián Agüero, it is him that I thank in the first place. In addition, he is one of the most analytical minds I know, and helped me to think things through more properly. I do admire his philosophical passion. Lorena Ramírez's comments ever astonished me for their wittiness: she knows how to put you in a tight spot. The

same can be said about Diego Papayannis: I was always a little afraid of discussing my work with him, due to his notorious argumentative skills. I must also thank Marcos Andrade, Cristina Astier and Mariona Rosell for taking their time to read my drafts, even though they had their own dissertation to write and my topic has little to do with theirs. Though not an LPM member, I cannot but remember here Biel Company. We started this journey together and have shared some wonderful moments.

Three other LPM members deserve to be thanked separately. Alberto Carrio and Laura Roth were my office mates until recently and for as long as I can remember. They learned to put up with me and my extemporaneous rants about personal identity, as well as with my more-than-worrisome obsession with the work of Derek Parfit. Finally, Cátia Faria. Over time, we became aware that we shared many things in common, not the least of these being our conception of practical reason and our stance in animal ethics. More than any other person she taught me how to do what is the best from a truly impartial point of view. For that, I shall never be able to thank her enough.

All these people contributed to what good there is in my work. Still, it would have been much harder to produce anything at all without the support of my family and close ones. My parents, Dolores and José, always encouraged me to pursue whatever career I preferred the most. In fact, it was my mother who first made me realise that what I really wanted is to be an academic, not a judge, as I had mistakenly fancied since I was a teenager. As to Sergi, my creative little brother, it is delightful to have him around –chatting with him always manages to put me in high spirits. I know that because of my work in this dissertation I failed to spend with them all the time love requires. They suffered for it. I only wish I could have known how to do it better.

I would now thank the person who has been closest to me during all these years, who has cherished me and supported me even to the point of inflicting severe harm to her happiness –Marina. No one has more caring traits of character, nor does more deeply sympathise with the joy and suffering of all

sentient creatures. I have tried to learn from her example. I hope I rendered all her sacrifices worth making.

To conclude, I feel the need to express my gratitude to several people who, although enjoying a different kind of existence than those previously mentioned, I nevertheless consider to be my moral heroes. Alonso, Pierre, Jean, Kostia, Kitty –you have my thanks for being beacons of moral integrity.

I benefited from an FI scholarship granted by the AGAUR during my Master course and the first three years as a doctoral candidate. With any luck, this institution will not find its money ill spent.

To all of them, again, my thanks.

Abstract

In this dissertation it is asked whether we ought to conceive of abortion, morally speaking, as killing someone or, rather, as preventing someone from existing. First, some preliminary distinctions about values and practical reasons are drawn. Secondly, the merits of different arguments for the position that abortion is like killing someone are compared. It is then proposed that the best such argument must consider that the relevant reason-giving facts pertaining to the morality of killing are those about the value the victim's life has for her and that diachronic identity is what matters in survival. Finally, a number of challenges to the ontological and normative assumptions of that argument are presented, it being ultimately found lacking. From the consideration of those objections it is concluded that abortion ought to be conceived, from the standpoint of practical reason, as deciding which of several possible beings shall become actual, and thus as a parcel of population ethics.

Resumen

En esta tesis me pregunto si, moralmente hablando, debemos concebir el aborto como matar a alguien o, más bien, como impedir que alguien exista. En primer lugar, presento algunas distinciones preliminares acerca de valores y de razones prácticas. En segundo lugar, comparo la solidez de diferentes argumentos a favor de la posición de que abortar es como matar a alguien. Propongo seguidamente que el mejor de tales argumentos es aquél que considera que, respecto a la ética de matar, los hechos relevantes que nos dan razones son aquéllos acerca del valor que la vida de la víctima tiene para ésta, así como que la identidad diacrónica es lo que importa en la supervivencia. Finalmente, presento una serie de desafíos a las asunciones ontológicas y normativas de dicho argumento, que al cabo encuentro deficiente. De la consideración de esas objeciones concluyo que es mejor concebir el aborto, desde la razón práctica, como una decisión acerca de cuál de varios seres posibles devendrá actual, y, por ende, como parte de la ética poblacional.

Preface

§1. *On the aim of this thesis*.

According to the latest World Health Organization (WHO) report, almost forty-five million abortions were performed in 2008, the last year for which such data are available[2]. Furthermore, the WHO estimates that each year there are an astonishingly eighty-five million unintended pregnancies[3]. Thus, yearly over half of all unplanned pregnancies terminate by induced abortion. Because the quantity of abortions and of women who consider whether to have one is so high, it is worth asking what we ought to believe about what we have most reason to do when confronted with the decision whether to abort a foetus[4]. The aim of this thesis is to search for an answer to this question.

This is not to be confused with the question of whether our social and political institutions ought to be designed so as to motivate women for or against abortion in certain circumstances. These are distinct questions. It is possible that, after reflection, we reach the conviction that in many occasions our institutions ought to motivate women to do what they have most reason to, be it killing the foetus or carrying the pregnancy to term. Yet it is also possible that

[2] WHO (2012, p. 19). I calculated this figure out of two others. The first is that that the absolute number of unsafe abortions in 2008 was estimated at 22 million. The second is that the proportion of all abortions that are unsafe was 29% in that same year. Hence, the absolute number (100%) of abortions, both safe and unsafe, in 2008 must have been roughly 44,900 million.

[3] WHO, *ibid*.

[4] As scientific categories go, it is not only the *foetus* I am concerned about, but also the *zygote* and the *embryo*. These three terms refer to different stages of development of prenatal human life. Of course, these stages, in themselves, carry no moral significance. For the sake of simplicity, I shall use the term foetus to refer to prenatal human life without distinction, as it is also usual in the literature. However, I shall be more precise whenever I want to draw attention to any specific developmental stage. I examine these stages in greater detail in APPENDIX A.

we come to believe that most of the time we ought to let women choose free from any sort of institutionally provided positive or negative incentive. There are, of course, many other possible answers between these two. Also, one may argue for an alternative of the last sort in several different ways. Perhaps the truth is that, although it is good that each woman decides about her pregnancy as she has most reason to do, what is best in our actual circumstances is that most women choose freely. On the other hand, it might be that on the best conception of our political relation with our fellow citizens there are values, and the reasons with which they provide us, which ought not to inform institutional design. Maybe the facts about abortion give us reasons of this sort[5]. In any case, an answer to this second, political question cannot be complete without an adequate answer to the first, non-political one. Even if it is believed that the political question is the one with *ultimate* importance, it must be conceded that the issue with which I shall deal possesses at least *instrumental* importance. This is one way how it is justified to treat it.

The facts about abortion which philosophers discuss are of two main types. Some are facts about the ontology of beings like us. In this respect, it is asked whether foetuses belong to the same kind as we do. Others are facts about value. Here it is asked whether the ontological question has any normative significance for –and, more generally, what sorts of considerations impinge upon– the morality of killing beings like us and beings like the foetus. I will not be defending a novel position in personal ontology. Nor will I introduce a new theory about the morality of killing. Whatever originality might be in this work, it must be found elsewhere. For what I purport to do is to take a broad outlook of the literature and to identify such consensus among the most plausible views as

[5] The contention that what may permissibly enter into the political justification of at least some issues of fundamental importance –such as the right to abortion– is not the whole normative truth, but only a part of it suitably identified, was famously proposed by John Rawls (2005 [1993]). A succinct criticism of Rawls's view on what he named *public reason*, and a defence of the contrary position that the whole truth about value must be brought to bear in public decision-making can be found in Ronald Dworkin (2006). In addition, the interested reader may find in Samuel Freeman (2004) a forceful reply to Dworkin's criticism.

might give us confidence in a judgement about what to believe about abortion. On the assumption that fulfilling this purpose might contribute to the academic debate, this is another way how my treatment here of this topic is justified.

Allow me a brief terminological point before explaining the layout of the thesis. Laypersons, including philosophers, usually mean something different by 'abortion' than physicians do. In the medical literature it is a term of art referring to the expulsion of the foetus from the womb before viability, that is, before it can survive outside of the uterine environment. The term 'abortion' is employed irrespective of whether such expulsion is natural or induced[6]. Moral philosophers, however, tend to use the term to designate the death of a foetus brought about by causing the aforesaid expulsion. Since it is its established usage in the literature, I will conform myself to it. In this way, whenever I say 'abortion' I shall mean 'the death of the foetus caused by its being removed from the womb'. Nevertheless, I am of the opinion that some clarity of exposition will be gained if 'abortion' in this sense is distinguished from 'foeticide', by which I refer to the death of a foetus caused by any means whatsoever[7]. From now on, 'foeticide' shall be my term of preference.

§2. *On the strategy that shall be pursued to achieve that aim*.

Some views, which I shall call *contrary to foeticide*, believe that this practice is a species of homicide, i.e., an instance of killing one of us. Nevertheless, I have come to believe that the most plausible views about foeticide are some of those which reject this claim. These views, which I shall call *favourable to foeticide*, conceive of it as rather more akin to choosing what possible beings shall

[6] See the entry 'abortion' in, for example, the Miller-Keane Encyclopedia and Dictionary of Medicine, Nursing, and Allied Health (2003), Steadman's Medical Dictionary (2006) or Mosby's Medical Dictonary (2009).

[7] Any of the works on abortion in the bibliography would serve as an example of the philosophical use of 'abortion'. I first observed the use of the term 'foeticide' in Hodge (1869).

become actual and thereby populate the world. My task is then, first, to show that these latter are indeed the most plausible views about foeticide. And, second, that this way of conceiving about foeticide really follows from the resemblances among them, in spite of their discrepancies.

Besides this PREFACE, the thesis consists of four chapters and two appendices. The chapters present the main argument for the conclusion, as well as the conclusion itself. The appendices discuss ancillary issues, such as the beginnings of human organisms (APPENDIX A) and the onset of foetal consciousness (APPENDIX B). The resolution of these questions is necessary in order to identify additional facts, both metaphysical and empirical, which condition the evolution of our reasons for and against foeticide throughout the pregnancy.

In CHAPTER I the better part of the conceptual framework which will be employed in subsequent sections is developed. There I distinguish, first, between two different conceptions, objective and subjective, of practical reasons. Later, I distinguish between, on the one hand, agent-neutral and agent-relative values (and the reasons they provide us with), and, on the other, personal and impersonal values. With these distinctions in our toolbox we shall be able to analyse the different views about foeticide discussed in other parts of the thesis. In this respect, I start in CHAPTER II by identifying the various strengths and weaknesses of different views contrary to foeticide. At the end of the chapter I expect to have successfully argued that the best of these is the one implied by the *Future-of-Value Account*, as it is called in the literature. This account is the object of sustained criticism from diverse fronts in CHAPTER III. There it will be observed how, even though standing on different, incompatible assumptions, the most plausible views against foeticide have much in common regarding how they conceive of the problem of foeticide. Finally, in CHAPTER IV, a brief summary of the preceding chapters is provided. As a conclusion, it is remarked how what we ought to believe is that deciding whether or not to kill a foetus is rather like deciding which of several possible beings ought to become actual.

CONTENTS

Acknowledgments vii

Abstract xi

Preface xiii

§1. *On the aim of this thesis*. xiii

§2. *On the strategy that shall be pursued to achieve that aim*. xv

CHAPTER I - VALUES AND THEIR REASONS 1

§1. *On the aim of this chapter*. 1

§2. *On the distinction between an objective and a subjective conception of practical reasons.* 1

§3. *On the distinction between agent-neutral and agent-relative reasons*. 7

§4. *On the distinction between personal and impersonal reasons*. 13

§5. *Concluding remarks*. 23

CHAPTER II - VIEWS CONTRARY TO FOETICIDE 25

§1. *On the aim of this chapter.* 25

§2. *On Anthropocentrism: preliminary remarks.* 27

§3. *On the impersonal version of Anthropocentrism.* 32

§4. *On the personal version of Anthropocentrism.* 46

§5. *On agent-relative Anthropocentrism. Some final remarks on this family of views as a whole.* 59

§6. *On the personal value of the contents of a life*. 62

§7. *Concluding remarks*. 69

CHAPTER III - VIEWS FAVOURABLE TO FOETICIDE 71

§1. *On the aim of this chapter.* 71

§2. *On denying that we are essentially human animals*. 74

§3. *On denying that identity is what matters*. 99

§4. *On claiming that a desire to live is necessary for someone's death to be prudentially bad.* 119

§5. *On denying that killing us is seriously wrong merely because it causes a prudentially bad death.* 132

§6. *On the permissibility of infanticide as a decisive objection to the permissibility of foeticide*. 141
CHAPTER IV - CONCLUSION 155
§1. *On the aim of this chapter.* 155
§2. *On the relatively best view contrary to foeticide* 155
§3. *On the more plausible views about foeticide*. 157
§4. *On what we ought to believe about foeticide.* 163
APPENDIX A - ON WHETHER AND WHEN A FOETUS BECOMES A HUMAN ANIMAL 165
APPENDIX B - ON WHETHER AND WHEN A FOETUS BECOMES CONSCIOUS 183
Bibliography 193

CHAPTER I
VALUES AND THEIR REASONS

§1. *On the aim of this chapter*.

In this chapter I shall comment on several formal aspects of practical reason. First, I shall distinguish between two incompatible conceptions about the nature of such reasons, namely, *Objectivism* and *Subjectivism*. Secondly, I shall introduce four types of practical reasons. Two of these stem from the distinction between agent-neutral and agent-relative values. The final two arise out of the difference between personal and impersonal values. All these categories will be of use in the two following chapters of the thesis, where I analyse the different arguments object of scrutiny.

§2. *On the distinction between an objective and a subjective conception of practical reasons.*

To solve a practical problem is to choose one of the competing courses of action or outcomes it presents us with as the most justified to carry out or bring about. In order to do that we need to know the reasons we have for and against, ideally, each of said courses of actions or outcomes. Among other things, this presupposes we have an account of practical reasons –what it is to have a reason, what kinds of them there are and how they differ from and relate to each other.

The topic of this book is not practical reason itself, but the specific practical problem of foeticide. Thus, I do not have the space to provide the reader with a

critical survey of the discussion about this topic. What I shall do instead is to present the distinction between a *Subjectivist* and an *Objectivist* conception of practical reasons. To that end, I will employ Derek Parfit's views about this issue, as lately developed[8]. This will be the basic constituent stuff of my arguments and I am perfectly aware that their cogency will therefore depend to an extent on the correctness of Parfit's views.

Every system of thought has its primitive concepts, that is, concepts that cannot be defined. According to Parfit, such is the concept of a *reason*. This doesn't mean that it is unexplainable, but only that it cannot be "helpfully explained merely by using words", but "by getting people to think thoughts that use [it][9]". One such thought would be that facts give us reasons "when they count in favour of our having some attitude, or our acting in some way". Another such thoughts would be that we always have a reason to avoid being in agony, or that the fact that a reliable weatherman has forecasted rain for tomorrow is a reason to believe that it will rain tomorrow.

As it can be seen, the concept of a reason is wider than that of a *practical* reason. Sometimes we have reasons for or against having some particular belief. These are usually called *epistemic* reasons. Practical reasons, for their part, come in two kinds. We can have reasons to want something to happen, i.e., reasons for having a particular desire. We can also have reasons to try and make that thing happen, i.e., reasons to act in some way[10]. One and the same fact, however, gives us both kinds of reasons, so that necessarily, whenever we

[8] That is, as it appears in parts of Parfit's *On What Matters* (2011a), especially pp.29-174. I will be dealing here with *normative* reasons, i.e., those that justify our conduct, which may or may not coincide with our *motivating* reasons, i.e., those that help to explain our conduct. The first may not coincide with the latter because we can act for reasons we believe we have and those beliefs might be false. See Parfit, *ibid.*, p.37, but also Raz (1975, pp. 18-19).

[9] Parfit, *ibid.*, p.31, for all the quotes in this paragraph.

[10] Parfit, *ibid.*, p.47.

have a reason to desire that the world be in such and such way, the content of our desire becomes also a rational aim of action[11].

I claimed that facts give us reasons and it is Parfit's opinion that the main theories about reasons agree on that –reasons are provided by facts[12]. The major disagreement is about what kinds of facts are reason-giving. Whereas Parfit believes that reasons are *object-given* and *value-based*, some others believe that they are *subject-given* and *desire-based*. The latter are the so-called *Subjectivist* theories about reasons[13]. They claim that reasons are subject-given in that they are provided by facts about the rational agent herself; and desire-based in that these are facts about what would fulfil the agent's present aims or desires. Of course, Subjectivism can come in different flavours: some may believe that all our actual desires are reason-giving, whereas others are more demanding and require that the relevant desires or aims be counterfactual –those the agent would have if he was ideally rational. Though I will not be endorsing here this particular conception of reasons, some authors that I will be discussing do, so it will be useful to keep it in mind[14].

---

[11] Notice that this is true even when it is in fact impossible for a rational agent to act in any way that might further the aim of satisfying her desire. For suppose that the fact that copious rain is a necessary condition for the harvest to be good provides the members of a famished community with reasons to desire that there be copious rain. Assume that this community has no access to any means by which rain may be caused to happen. The same fact still gives them a reason to try and make it rain. Such reason for acting does not cease to exist merely because the agents cannot act upon it, just as it does not come into existence once they can. But because this is a reason which these agents cannot act upon, they cannot be criticised for failing to do so. That would still be so even if to try and make it rain was, using terminology which will be explained below, what these agents have most reason to do. I believe this is all that it takes for this account of practical reasons not to run afoul of the doctrine that 'ought implies can'.

[12] For a similar view on the relation between facts and reasons see Raz, *ibid.*, p.17 and Raz (2002, pp. 22, n.4).

[13] Parfit's (2011a) first presentation of this distinction is in pp.45-47, and one of his main goals throughout the two volumes of *On What Matters* is to debunk it. But see Parfit (1984, pp. 120-126) for an early discussion about this topic.

[14] See CH. III, **§4**.

Now, subjectivists agree that we have reasons to have *some* desires. The fact that the obtainment of some state of affairs is necessary for the fulfilment of some desire or aim gives us an *instrumental* reason to try and make such state of affairs obtain. Also, we have the corresponding instrumental desire that it obtains. This is all perfectly consistent since our reasons to have an instrumental desire are still both subject-given and desire-based, so that none of the tenets of Subjectivism has been broken. What of the desires towards whose fulfilment some other desires are instrumental? Do we have reasons to have them or not? That depends on whether these desires are also instrumental. If they are, we do have such reasons. But, of course, the buck must stop somewhere –there must be some non-instrumental (*telic*) desires and aims that start the chain of desires and from which all our instrumental reasons derive their force. However, all reasons are desire-based and there are no further desires or aims beyond telic desires and aims. Subjectivists must conclude that there are no reasons that can justify our non-instrumental desires.

So, what does Parfit's *Objectivism* about reasons consist in? We have reasons to have desires and aims, and to do what might achieve them, and these reasons are given by facts about the *object* of those desires. This is the fact that the object of the desire is either worth producing or worth preventing, that is, *valuable* in some way. Here the distinction between telic and instrumental desires applies too, albeit with a difference –the Objectivist must claim that we not only have instrumental practical reasons, but also that some events are *valuable as ends*, not merely as means to achieve some further event. Thus, some things, those worth for their own sake, provide us with telic reasons and telic desires[15]. We may have other reasons to want other events to happen, but just because they contribute to bring about something valuable as an end.

[15] Parfit (2011a, p. 52).

What kinds of things are valuable as ends? Well, Parfit gives us some examples[16]:

Being in agony.

Feeling less lonely.

Seeing the sublime view from the summit of some mountain.

Understanding how life or the Universe began.

As these examples show, some things are positively valuable (good), whereas others –like being in agony- are negatively valuable (bad)[17]. Arguably, among valuable events, some are more valuable than others so that they give us stronger reasons to prefer their happening to that of less valuable events, which give us weaker practical reasons. Negative value, or disvalue, behaves in a similar fashion, although providing us with reasons not to prefer the happening of some event. In order to know what reasons we have for or against trying to make some event happen we will need to combine all the reasons given by the different facts about the valuable attributes or nature of such event[18].

Thus, whenever we face a practical problem, we have to assess all the relevant reason-giving facts about the several possible ways we may act[19]. All things

[16] The first example appears in page 31 and 57 of Parfit (2011a), and later prominently during the course of the Agony Argument (2011a, pp. 73-82). For the others see Parfit, *ibid.*, p.52.

[17] This doesn't correspond to Parfit's usage of 'good' and 'bad', which implies that we have *strong* reasons for or against something, see Parfit, *ibid.*, p.38.

[18] The keen reader will notice that sometimes I will predicate value of facts, some other times of events or states of affairs and even others of attributes. Certainly, if an object of our desire is to count as valuable, it must be because some attribute or attributes of it. If that is the case, I see no reason not to say that such attribute is valuable. "These are merely different ways of making the same claims", as Parfit (2011b, p. 743) says whilst discussing a similar point.

[19] Or about the several desires we may have. All I say here is true of desires too.

considered, we may have either *decisive* reasons or *sufficient* reasons for or against a course of action[20]:

> *Decisive reasons*. We have decisive reasons to act in some way just in case the reasons we have for acting in this way are *stronger* than those we have against acting in any other way.
>
> *Sufficient reasons*. We have sufficient reasons to act in some way just in case the reasons we have for acting in this way are *not weaker* than those we have against acting in any other way[21].

Of course, what we ought to do is what we have decisive reasons to do, which includes doing *one* of the things we have sufficient reasons to do[22].

There is one further point worth making. What we have most reason to do need not coincide with what we believe we ought to do[23]. Reasons are given by facts, and we may well be ignorant or uncertain about the relevant facts. The reasons we believed we had can be important in order to judge the rationality or irrationality of our desires and actions, that is, to judge whether our conduct is praiseworthy or blameworthy. It must be recognized, however, that someone can, at the same time, act rationally *and* fail to do what she had decisive reasons to. One may act according to the reasons one would have had if one's

[20] These definitions are verbatim transcriptions of Parfit (2011a, pp. 32-33).

[21] I may use sometimes the metaphor of *weight* instead of the metaphor of *strength*, and claim that some reasons outweigh others or are outweighed by them. Nothing different will be meant in such occasions, being only for the sake of style, and both metaphors being current in the literature. Also, I will be using the phrase '*x* has most reason to *φ*' as equivalent to '*x* has decisive reasons to *φ* '.

[22] For, by definition, whenever we have sufficient reasons to *φ* we also have sufficient reasons not to *φ*. Since these are our only alternatives of action, we can rightly infer that what we have most reason (i.e., decisive reasons) to do is either to *φ* or not to *φ*.

[23] Parfit, *ibid.*, pp.33-37.

beliefs were true, but those beliefs might unfortunately be false. We may as well call upon Parfit for an example to illustrate this distinction:

> "Suppose that, while walking in some desert, you have disturbed and angered a poisonous snake. You believe that, to save your life, you must run away. In fact you must stand still, since this snake will attack only moving targets. Given your false belief, it would be irrational for you to stand still. You ought rationally to run away. But that is *not* what you ought to do in the decisive reason-implying sense. You have no reason to run away, and a decisive reason *not* to run away."[24]

In my discussions, unless stated otherwise, I will assume that we know all the relevant facts about the foetus. According to the framework I have sketched in this section, the disagreement among the different philosophers that have dealt with the moral problem of foeticide will stem from a disagreement about which are the relevant reason-giving facts, or about the strength of the reasons they give us, or both.

§3. *On the distinction between agent-neutral and agent-relative reasons*.

In this way, during the course of this dissertation we will be reviewing the various strengths and weaknesses of competing arguments about the normative status of foeticide. My analysis of these arguments will hinge on most occasions on several distinctions about the kind of reasons for acting they presuppose, either explicitly or implicitly. In the previous section I explained the distinction between sufficient and decisive reasons for a course of action, all things considered. However, reasons not only differ in their *strength*, but also in

[24] Parfit (2011a, p. 34). The italics are Parfit's. Another example and an early treatment of the *having a reason*/*being rational* distinction can be found in Parfit (1984, p. 25).

their *structure*. In the terminology about the values on which reasons are based, we could say that things can be valuable in several different ways[25].

Let us picture Clara. She is a very inquisitive young girl, but has the regrettable flaw of character of lying whenever she believes it will advance her interests. After a series of particularly blatant lies, her father, Peter, decides to intervene and lecture her. He succinctly puts it to her daughter – 'You ought not to lie'. For our purposes here, we may as well construe Peter's claim as implying

(1) You have a reason not to lie.

Clara, being as I said a very inquisitive girl, is not quite satisfied with this. She needs to know what is so bad about lying. After all, she argues, if you have a reason not do something, then it must be because there is something bad about it. 'So', she asks, 'why is it bad if I lie?'.

Peter, of course, is baffled, both by the acumen of the question and the fact that it was posed by his little girl. Well, what is bad about her lying? After a while, he realises there are two possible rough answers to that question:

(2) It is bad that *Clara* lies.

(3) It is bad that lies are told.

So, imagine Peter makes up his mind and chooses (2) as the most plausible description of what is bad when her daughter lies. Generalizing a little bit, and being philosophically inclined, he arrives to a more precise account of what is bad with anybody's lying, and in what way:

[25] In this section I follow Nagel (1970, pp. 91-98) and Parfit (1984, pp. 27, 55, 93, 102-104). Notice, however, that Nagel calls what here I will be referring to as agent-neutral reasons 'objective' and what here I will be referring to as agent-relative reasons, 'subjective'.

(4) It is bad that there is some person, *p*, and some state of affairs, *s*, such that *s* entails *p*'s lying,

and

It is bad in that it gives *p* a reason to try and prevent *s* from obtaining.

Armed with his little moral theory about lying, Peter can now try and give Clara a more detailed reply to her question. 'Look', he might say, 'it is a bad thing when you, Clara, or anybody else for that matter, lies. Because of the way it is bad, each of us has a reason not to tell *their* lies, and each has to exercise herself to take that reason into account and act according to it. It may come a point when you realize that you have the opportunity to make people lie less by your telling some lie. Well, even in that case, you have a reason not to lie, because you never had a reason to *minimize* the number of liars or lies in the first place, but a reason not to tell *your* lies. That is true of everyone'.

At first, Clara seems quite satisfied with this answer, but after a while, she begins to have her doubts. What if by telling a lie she could prevent someone from telling an even worse one? And wouldn't it be great if there were fewer liars or fewer lies? That appears to be a good thing to Clara, which might give her reasons to lie at least as strong as her reason not to tell lies. All this she declares to her father. Sometime later, Peter realises that perhaps he made the wrong decision and that instead of (2) as an answer to Clara's original question, he should have chosen

(3) It is bad that lies are told.

This Peter neatly puts in a more formal way as:

(5) It is bad that there is some person, *p*, and some state of affairs, *s*, such that *s* entails *p*'s lying,

and

It is bad in that it gives all rational agents a reason to try and prevent *s* from obtaining.

Of course, this means Peter can give Clara a rather different answer to her questions. 'It is a bad thing when people tell lies, Clara. This means, sure, that you have a reason not to lie. But do not forget that what matters here is that there are as few people telling lies as possible. So, sometimes, though you may have a reason not to lie, you may also have a stronger reason to do so, because, all in all, that is what would cause there to be fewer people lying in the world. That is true of everyone'[26].

I do not know whether (4) or (5) is the best little moral theory about what is bad about lying and in what way it is bad, nor am I presently interested in that. The point of the story is to bring forth two senses in which things, or states of affairs, can be good or bad in a reason-implying way. Sometimes we individuate a state of affairs and judge that, somehow, it is its *mere obtaining* which is good or bad –as in (5)–, so that we all have a common reason to try and cause it to happen, or to try and prevent it from happening. When things are good or bad in this

[26] Strictly speaking, (5) is just *one* possible way of rendering (3). An alternative way of doing so would be to do without a variable ranging over persons altogether:

(6) It is bad that there is some state of affairs, *s*, such that *s* entails that there are lies.

and

It is bad in that it gives all rational agents a reason to try and prevent *s* from obtaining.

In a world where there could be lies caused otherwise than by people telling them, (6) would give us additional reasons for acting (5) would fail to provide. As it is, though, in the actual world (5) and (6) provide us with the same reasons.

way, we say that they are good or bad in an *agent-neutral* way and they give us agent-neutral practical reasons:

> *Agent-neutral valuation*: a state of affairs is valuable in an agent-neutral way just in case
>
> (a) It gives *all* rational agents a reason for acting
>
> and
>
> (b) It assigns as the *common* rational aim for those agents the obtainment of such state of affairs.

The notion of agent-neutral values and reasons will be pretty useful in the coming analysis of philosophical arguments for and against foeticide. Just consider some of the things you may remember about everyday discussions about this issue in the media or in politics. When your average pro-lifer insists that abortion is tantamount to the murdering of innocents because the foetus is a human being, and human life is sacred, she might just be singling out an attribute of the foetus (*being human*) and saying that it is valuable in an agent-neutral way, and very much so. On her view, we all have a strong reason not to kill the foetus. Something similar can be said about the typical pro-choicer who argues back by pointing out that the foetus is not a person. Here the assumption is that it is the attribute of belonging to the kind *person* (however it may be defined) that gives us a reason not to kill each other. Whether we have reasons or not against killing the foetus, they must be found elsewhere. Certainly, the actual philosophical picture is much more sophisticated than that, and there are many attributes beyond those two (*being human* and *being a person*) that have been put forward as candidates for what gives us all our main reasons against foeticide. Nevertheless, this provides us with a taste of things to come.

Regarding agent-relative valuation, consider the case of *prudence*. Prudence is a central case of this second way in which a state of affairs can be good: it is good that *each* of us strives to make *her* life as flourishing as possible[27]. When we say that it would be prudent for someone to spend his time reading *Lolita* instead of playing video-games we are probably appealing to something like,

(7) It is good that there is some person, *p*, and some state of affairs, *s*, such that *s* entails that *p*'s life is maximally flourishing,

and

It is good in that it gives *p* a reason to try and cause *s* to obtain.

Here, we imply nothing about the existence of reasons to contribute to the flourishing of lives in general. Even if such reasons exist and are common to all agents, they cannot be derived from (6). We are only allowed to claim that *each* of us has a reason to care for the flourishing of *her own* life. If that were our supreme rational aim, it would be irrational for us to help others flourish when that would, all things considered, undermine our efforts to achieve it.

When things are good or bad in this way, we say that they are good or bad in an *agent-relative way* and they give us agent-relative practical reasons:

Agent-relative valuation: a state of affairs is valuable in an agent-relative way just in case

(a) It gives reasons for acting *only* to those rational agents who also figure as one of the terms in the relation featured in the description of the state of affairs,

and

[27] For prudential or self-interested rationality as a paradigm of agent-relativity see, for example, Parfit (1984, pp. 3-5 and 55).

(b) It restricts *each* agent's rational aim to the obtainment of the state of affairs in which *she* so figures.

Most of the arguments that shall be reviewed in subsequent chapters can be construed as assuming that our reasons against-killing foetuses are agent-neutral. At least one, however, explicitly presents its case as grounded on agent-relative reasons.

§4. *On the distinction between personal and impersonal reasons*.

Now, important as the distinction between agent-neutral and agent-relative reasons is[28], there is a second dichotomy related to different ways values may give us reasons, which also reveals itself as a powerful tool for analysis[29]. Let us reprise Peter and Clara's story. We had left them discussing the badness of lying, and they seemed to have agreed that we all have a reason against it, including Clara. Still, she feels that there is more to know. Particularly, she wonders whether her lying is bad because it is bad for her or because it harms those who are told the lies. 'Or perhaps', she muses, 'it is just that lies are such an awful thing'.

As is now her custom, she seeks an answer from her father. Who is lying bad for? As it happened before, Peter must choose between two competing suggestions. Either,

[28] The reader may use the table at the end of this chapter for a quick reference-guide to the distinctions discussed in both the current and the previous section.

[29] In this section I follow Dworkin (1994, pp. 68-81), McMahan (2002, pp. 330-331), Boonin (2003, p. 42) and Kamm (2007, pp. 227-230).

(8) Lying is bad for someone

Or

(9) Lying is *just* bad.

Being it difficult for him to picture how something could be just bad without it being bad for anyone, Peter chooses (8) as the most plausible starting point for an account. ‘Look, Clara,’ he begins, ‘lying can harm both you and other people. That is why it is bad. It is pretty easy to see why you have a reason not to tell a lie when it can do harm to others, perhaps by withholding from them important information that could do them good, or when it is disrespectful towards them because they become the object of your mockery. But even if there is an instance when you could tell a lie which is completely innocuous in that regard, it could still be bad for you. You see, some people think it is good for you to be honest, because that is one aspect of being a decent person. This is true of everyone’. Being the meticulous person he is, Peter insists once more in presenting this account of the badness of lying in a more formal way:

(10) It is bad that there is some person, *p*, and some state of affairs, *s*, such that *s* entails *p*’s lying,

and

It is bad in so far it detracts from someone’s well-being.

This strikes Clara as a very persuasive explanation of why lying is bad –it gives you a reason not to do so when you can harm yourself or others, but not otherwise. As to the possibility that lying is just bad, she finally concludes that if lies seemed so awful to her it was, after all, only because of the possible harm they might inflict. The thought, however, that there could be things that are *just* bad or *just* good, lingers in her mind. As it happens, the girl loves trees and she

hears about the existence of a particularly majestic and ancient one. Could it be *just* good that this tree exists? After hearing all this, Peter is not so ready to discard the sort of value suggested by (9). He is not particularly fond of trees, but he realizes that works of art pose a similar problem.

The question is whether there are at least some things, like some works of nature or some works of art, whose existence makes the world a better place, even when the fact that they exist is good for no one. On a desire-based conception of reasons, this would make little sense, because whenever we have a reason to preserve a majestic tree or a particular work of art, it must be because their existence is the object of someone's instrumental or telic desire. Yet in an object-given, value-based view we must consider it at least *conceptually* possible. What would that mean? Roughly put, that we would have reason to prefer, *caeteris paribus*, a world with works of art to one devoid of any, even if we could never experience them. How so? First consider that even if on the best substantive account of the good life it turns out that works of art belong to the inventory of objective goods, the most we can say is that their *being experienced* is good for us. Since their *existence* is necessary for their being experienced, however, we can also say that it is instrumentally good, which gives us a reason not to destroy them. Thus, if we say that beyond being good for someone they are *just* good, we must be implying something else –that their mere existence, instrumentally good for us or not, is valuable in some further way[30].

Again, it is not my goal here to take a stance on the nature of the badness of lying, nor do I commit myself to the claim that works of nature or works of art are just good in the way described. I hope that these examples have been

[30] Cf. Kamm (2007, pp. 228-229) who, regarding the possibility that trees or works of art have moral status though lacking a good of their own, comments: 'independent of valuing and seeking the pleasure or enlightenment it can cause in people, a thing of aesthetic value gives us (I think) reason not to destroy it. [...] I save it for no other reason than that it will continue to exist'.

helpful as an introduction to the third and fourth way in which things, or states of affairs, can be said to be good or bad in a reason-implying way. Sometimes things are good or bad for people. What matters to us in these occasions is that someone's well-being is affected positively or negatively. Whenever this is the case, we say that some state of affairs is *personally valuable* and that it gives us *personal* practical reasons:

> *Personal valuation*: a state of affairs, *s*, is personally valuable just in case that there is an *s* and an entity, *x*, and the predicate '*s* is valuable for *x'* is satisfied by both.

That being said, I believe two clarifications are in order. First, don't let the adjective 'personal' fool you. There is no conceptual, necessary link between personal values and what is good or bad for persons. That is why I predicate the relation 'being valuable for', and its variants, to an entity, and not to a rational agent. Beyond beings like you and me –which is the rough, preliminary meaning I will give to 'person'- there are many entities with a well-being of their own of which the aforesaid relation can be sensibly predicated. Many non-human animals are a fine example of this. I know of quite a few things that make my dog happy, and some that would make him miserable. I have reasons to pet him and keep him company, and sometimes those reasons may be decisive[31].

Secondly, it would be a mistake to believe that whenever something is valuable in a personal way, it must *necessarily* be valuable also in an agent-relative way. Indeed, it can be so *contingently*. This is the case with *self-interested* reasons. Recall our discussion about prudential considerations. If someone would be better off reading *Lolita* rather than playing video games, then he has an agent-

[31] So perhaps it would have been more transparent to name this kind of values, and their reasons, *sentient-affecting*. In this way, it would have been clearer that they have to do with what is good or bad for all sentient (or conscious) beings. But I found the phrase too cumbersome, and the usage of the expression 'personal' or 'person-affecting reason' too established in the discussion, too attempt the change.

relative reason to do so, and since it is given by facts about his well-being, it is also a personal reason.

Yet states of affairs that are personally valuable can provide us with *agent-neutral* reasons. In fact, that is the sort of value that many pro-lifers attach to foetal live. Just as some prefer to stress the fact that foetuses are human beings, some find it more persuasive to point out that, at least in developed societies, unless it is killed, the average foetus will have a life worth living. Thus, they appeal to facts about the foetus's future well-being and claim that they provide us all with a reason to let it live. It is precisely *because* its death would be bad for the foetus that we ought not to kill it. Pro-choicers, of course, will argue that it is false that death can be bad for the foetus, or at any rate that it is not so bad as to give us decisive agent-neutral reasons against ending its life.

Let us now turn again to the fourth way something can be good or bad. Sometimes outcomes are *just* good or bad, even if they are good or bad for no-one. What matters to us in these situations is sometimes captured in phrases like 'the world becomes a better place' or 'a worse place' –we are concerned with the overall value of the actual world. Whenever something is valuable in this way, we say that it is *impersonally valuable* and that it gives us *impersonal* practical reasons:

> *Impersonal valuation*: a state of affairs, *s*, is impersonally valuable just in case the predicate '*s* is valuable ***simpliciter'*** is satisfied by it.

When we have decisive personal reasons to benefit someone but we act against them, it is not only true that we have acted wrongly, but we also say that we have wronged the individual to whom the benefit was due. When we act

against decisive impersonal reasons, though it is still true that we have done wrong, it cannot be true that we have wronged anyone[32].

There are, again, a couple of conceptual points worth making in order to avoid confusion regarding the notion of impersonal values. First of all, just as personal values are not necessarily agent-relative, impersonal values need not be agent-neutral. Just suppose someone obsessed with etiquette conceives of the value of dressing according to the appropriate standards as impersonal and agent-relative, and claims something like 'each person has a reason to dress according to etiquette, even when it is against her overall well-being'. Certainly, this is a ludicrous normative claim, but here we are not concerned about its plausibility, but about its formal structure:

(11) It is good simpliciter that there is some person, *p*, and some state of affairs, *s*, such as that *s* entails that *p* dresses according to etiquette,

and

It is good in that it gives *p* a reason to try and cause *s* to obtain.

What is valuable here has nothing to do with what is good for *p*, but simply with his dressing according to etiquette. As per our definition, that makes *s* impersonally valuable. Also, it is not valuable in that we all have a reason that *p* dresses properly, not even a reason to ensure that everybody dresses properly. Our rational aim is restricted in the typical agent-relative way, namely, each of us has a reason to care only about her own dressing.

Secondly, although it is not possible for the very same fact to provide us with both agent-neutral and agent-relative reasons, given the definitions, I believe

[32] See, for this particular point, McMahan (2002, p. 331) and Kamm (2007, p. 230). Perhaps the name of this kind of value, and of the reasons it gives us, ought to be changed in a way similar to the one I referred to in the previous footnote.

there is a way for a personally valuable entity to be also impersonally valuable. Consider

> That *s* obtains is valuable for *p*
> (e.g., that *p*'s achieving professional success obtains is good for *p*)

If we add the intermediate premise:

> That something personally valuable obtains is impersonally valuable.

Then we could conclude

> That *s* obtains is valuable *simpliciter*
> (e.g., that *p*'s achieving professional success obtains is good *simpliciter*)

In this way, we may say that there are two sorts of states of affairs which can bear impersonal value. Some are states of affairs the impersonal value of whose obtainment can be accounted for by aggregation of the personally valuable things that figure in it. In these cases, personal values are explanatorily prior to impersonal values. Others are states of affairs such as the existence of Clara's tree or of Peter's works of art the impersonal value of whose obtainment cannot be explained by appealing to a prior personal value.

Now, there are at least two objections that can be made against each of the sorts of impersonally valuable states of affairs that I have described –that they are incoherent or that they play no role in practical reason. A lot of work will be saved If these objections are successful, since then I would be able to level them against all those accounts of the normative significance of foetal life that render it impersonally valuable. As to the objection that they are incoherent, when I first introduced the examples of the majestic tree and Peter's works of art I already suggested that though the notion of impersonal values would hardly fit in a subject-given, desire-based account of practical reasons, in an

object-given account they are at least conceptually possible. Of course that doesn't mean that we should incorporate these values into our ontological (or, if you will, methodological) inventory –it must be shown that we need them to justify intuitively plausible decisions.

In order to do just that, consider the following situation,

> *The World Maker*. You are presented with a barren universe and are given a choice between two scenarios to create. In each one there exist three persons whose lives will be worth living. In the first scenario, let us call it *Good World*, your universe would be populated by individuals A, B and C, whose respective levels of overall well-being would be, if created, 60, 50 and 40. In the second scenario, *Mediocre World*, your universe would be populated by three different individuals, D, E and F, whose respective levels of well-being would be 30, 20, and 10.

Which world ought you to create?[33] I feel strongly inclined to choose *Good World*, but I will now try to explain why it is hardly possible to justify that choice by appealing to personal values. Imagine that at the time of the choice I find myself in a sadistic mood and decide to create *Mediocre World*. After a while, D, E, and F discover that I could have chosen differently and created *Good World*. They feel that I did wrong and they reproach me so. 'See', I might defend myself, 'you have no reason to criticise me, for I have done *you* no injury. You would not have been better off if I had created *Good World*, nor are you worse off because I created yours. True enough, if I had created the other world, each of its three inhabitants would have enjoyed a higher degree of well-being. Yet it

[33] Our World Maker is, of course, facing an instance of the Non-Identity Problem in Same Number Choices as described in Parfit (1984, Part Four), and previously in Parfit (1982). I am not worried, however, whether the way out of this dilemma, as suggested by Parfit in the works just referenced and again in (2011b, p. 744), is justified, though it is the one I will be using here. I am only interested in that it presupposes the existence of impersonal values. But see also, for instance, Glover (1977, pp. 66-69), Steinbock (1992, pp. 37-40; 2011, pp. 31-34) and Singer (2011, pp. 108-111).

is also true that those inhabitants would not be the three of you, but A, B and C. In fact, you would not exist at all. It was impossible for me *both* to create the other world *and* make you better off. Since anyway your lives are worth living, my actions are beyond reproach.'

Perhaps it will help to justify our preference for *Good World* if we describe each of the worlds in a different way. Since we are assuming

> That something personally valuable obtains is impersonally valuable.

we could say

> That anybody's life is personally good for her is impersonally good.

So that

> *Good World* is the world in which there exist three valuable states of affairs–the one of value 60, the one of value 50 and the one of value 40.

and that

> *Mediocre World* is the world in which there exist three valuable states of affairs–the one of value 30, the one of value 20 and the one of value 10.

What we have done is redescribe the circumstances in each world without reference to persons and without resorting to claims about what is good for anyone. Also we have assigned to the relevant states of affairs the same numerical value as the corresponding personal value of each person's life[34].

[34] These figures represent the weight of the impersonal value of each state of affairs relative to the weight of the impersonal value of any other state of affairs, and tell us nothing about their weight relative to personal values. It might well be that sometimes personal values give us stronger practical reasons than impersonal values.

This would be similar to the sort of value we conceived of for works of nature or works of art –we would have reasons to prefer the existence of the states of affairs of greater value, even if no one is made better-off by our choice or worse-off by making a state of affairs of lesser value to obtain. Thus, if we take the figures to stand for the strength of the reasons for the existence of each of these events and if we adopt an additive principle for aggregating value, we can easily conclude that, all things considered, we have decisive reasons to create *Good World* and, consequently, decisive reasons not to create *Mediocre World*. If we wanted to justify our decision in this imaginary situation, then, we could appeal to impersonal values and the reasons they give us, but not to personal values.

Even assuming that we have a sufficiently strong *prima facie* case for the necessity of impersonal values in practical reasoning, it could be objected that I have only managed to suggest how that something personally valuable obtains is impersonally valuable, but not how other states of affairs like the existence of a majestic tree or of a work of art could be so. I agree. But now we have descended onto a debate about *which* sorts of things can be impersonally valuable, not about the cogency of impersonal values themselves. This is a discussion that I leave for later on, when we discuss particular positions about the value of foetal life. Some of the accounts that seem to underlie prominent popular arguments against abortion can be plausibly construed as appealing to impersonal values. I have already discussed how some pro-lifers appeal to the fact that the foetus is a human being. I said that they are probably attaching some agent-neutral value to the attribute of belonging to the kind 'human being'. Furthermore, some of them believe that the value of human life is impersonal, which they try to convey by using the term 'sacred'[35]. This construal of their position is very helpful because it allows us to accommodate some other intuitions about human life that adherents to the so-called *sanctity of human life doctrine* usually share. Since impersonal value need not be concerned with

[35] As we shall see (CH. II, **§3**) 'sacred' is Dworkin's (1994, p. 70) label for a species of *impersonal* value, which he calls *intrinsic*.

facts about the well-being of persons it is perfectly coherent to say, of a particular situation, that, although there is someone the personal value of whose life is negative, the impersonal value of her life is tremendous and that, all in all, we have decisive reasons to keep her alive[36].

§5. *Concluding remarks*.

A moral theory about foeticide ought to provide us with an account of the normatively relevant facts, including an account of the value of foetal life. Ideally, such a theory ought to bear in mind and employ the distinctions between agent-neutral/agent-relative and personal/impersonal value that were here presented. The contemporary discussion about the morality of foeticide is, nevertheless, several decades old now. Some authors used these categories of value, but referred to them by different names. In many occasions the authors I will be discussing simply do not rely on them in their argumentation, maybe because they employed other taxonomies, maybe because they approached the issue from a lower level of abstraction. Whenever this happened, I set upon myself the task of construal and reconstruction along the lines of the different ways of being valuable that I have discussed.

[36] I shall review this doctrine in CH. II, **§2** through **§5**, under the name of *Anthropocentrism*.

Ways Something can Be Valuable[37]	*Personally* Being valuable for someone	*Impersonally* Being valuable *simpliciter*
Agent-neutrally All *p* are assigned a common rational aim	All *p* have the common aim to try and make *s* happen, because the obtainment of *s* is valuable for someone. Example: *We all have a reason to try and make it that lies are not told whenever someone would be harmed otherwise.*	All *p* have the common aim to try and make *s* happen, because the obtainment of *s* is valuable *simpliciter*. Example: *We all have a reason to try and make it that lies are not told because otherwise the world becomes a worse place.*
Agent-relatively Each *p* is assigned an aim regarding *s(p)*	Each *p* has the aim to try and make *s(p)* happen, because the obtainment of *s(p)* is valuable for someone. Example: *Each of us has a reason not to tell lies whenever someone would be harmed otherwise.*	Each *p* has the aim to try and make *s(p)* happen, because the obtainment of *s(p)* is valuable *simpliciter*. Example: *Each of us has a reason not to tell lies because otherwise the world becomes a worse place.*

[37] Take *s* as a variable ranging over states of affairs; *p* as a variable ranging over rational agents and *s(p)* as a variable ranging over those states of affairs in which *p* figures as one the terms of the relevant relation featured in the description of *s*.

CHAPTER II
VIEWS CONTRARY TO FOETICIDE

§1. *On the aim of this chapter.*

There are several possible ways of classifying the different views on the moral status of the foetus. One may want to stress how moral philosophers differ regarding what matters in these cases. Then one could use the distinctions I presented in the previous chapter (especially the distinction between personal and impersonal values), provide the best version or versions of each competing account of the value of foetal life and observe what can be inferred from such an analysis. However interesting that might be, this way of classifying the views of the moral philosophers that have had something to say about this issue does not correlate with the for/against divide regarding foeticide. I agree that a taxonomy of philosophical positions on an issue based on the authors' *assumptions* might usually be methodologically preferable and much more illuminating than a taxonomy based on the authors' *conclusions*. Yet I also believe that my purpose in this chapter –namely, identifying the best argument for a view contrary foeticide– warrants here an exception.

Now, by 'view contrary to foeticide' I do not mean just the view that we always have *decisive reasons* against foeticide. I am rather referring to all those views that claim that our reasons not to kill a foetus are, caeteris paribus, *at least as strong* as our reasons not to kill individuals like you or me, that is, adult human beings whose lives are presumably worth living. And I will refer to those views as contrary to foeticide even if they allow that sometimes we have sufficient or even decisive reasons to kill a foetus, just as, quite plausibly, we exceptionally have sometimes sufficient or even decisive reasons to kill innocent adult human beings. The main feature of the views contrary to foeticide as I have

characterised them is, then, the claim that the requirements for the justification of foeticide are at least as demanding as the requirements for the justification of killing individuals like you or me, whatever those requirements might be.

This main feature of views contrary to foeticide, true enough, makes the strength of the case against it dependant on the strength of the case against killing human adults in ordinary circumstances. Needless to say, the belief that killing people is very seriously wrong except under special circumstances is backed by unusually powerful intuitions, so a successful attack by 'views favourable to foeticide' cannot come from that front. What these other views claim is that the reasons we have against foeticide are, caeteris paribus, *much weaker* than the reasons we have against killing individuals like you or me and that, as a consequence, the requirements for the justification of killing a foetus are *much less* demanding. This might be because whatever facts give us reasons against foeticide are *not the same* as those that give us reasons against killing individuals like you or me or because, even if they are, the relevant reason-giving properties are present to a lesser extent.

But I shall not deal with these views favourable to foeticide until CHAPTER III. In what follows, I shall try to cast the contrary views in their best light. As I explained before, all these views start by articulating an account of the intuition that we usually have decisive reasons not to kill beings like you or me (hereafter for short an account of the 'wrongness of killing'), then apply that account to the case of foeticide and purport to show how we have similarly strong reasons not to kill a foetus. I will argue that the most robust among these contrary views are those that claim that our reasons not to kill people are given by facts about the *personal value* of their lives. Other accounts are, by comparison, too problematic. Let us see why.

§2. *On Anthropocentrism: preliminary remarks.*

The attribute traditionally identified as the relevant reason-giving feature against killing, shared both by human foetuses and typical human adults, is (unsurprisingly) *belonging to humankind*. Quite a few contemporary philosophers find this to be the most plausible account of the wrongness of killing[38] and it remains influential among large numbers of people. In fact this account of the wrongness of killing is but an entailment of a much general view that identifies the attribute of belonging to humankind as giving us decisive reasons for a fundamental equal treatment of all human beings, irrespective of their other individual attributes. I believe that this view is mistaken, but I do not purport to disparage those who defended it in the past or still adhere to it today. It was probably for the good of us all that the conviction that all humans are fundamentally equal became widespread, since it displaced mistaken views as to what constituted valid grounds –race, gender, ancestry, and so forth- for a fundamentally unequal treatment of groups of people. It was bad for other conscious beings, though. Because of their belonging to other kinds, it was deemed that our treatment of them was much less morally demanding also irrespective of their individual features. But we are now in a position to recognise the problems of this account and to propose alternatives to it that will do a better job of singling out what matters pertaining the morality of killing.

Let us try to understand better, though, what this *anthropocentric* view, as Jeff McMahan calls it, claims[39]. Anthropocentrism has been defended on religious grounds, and more relevantly to our cultural milieu, it is still one important part of Christian doctrine[40]. I will not consider here, though, possible arguments for

[38] Such as John Finnis, Christopher Kaczor, John Keown, Patrick Lee or Michael J. Perry (see appropriate references in subsequent footnotes).

[39] McMahan (2002, pp. 206-228).

[40] It is common knowledge that Christian morality is anthropocentric in McMahan's sense. Consider, for example, the following passage: "Life is *always* a good. […] Why is life always a good? […] Man has been given a sublime dignity, based on the intimate bond which unites him

this view based on religious beliefs. For one, it is beyond the scope of this work to enter into such considerations. Also, I seek a view contrary to foeticide about which it is plausible to suppose that we have sufficient reasons to endorse it, whether we are religious or not[41]. Finally, there are some openly Christian authors which have made the effort of offering grounds for their convictions regarding foeticide in terms of pure secular philosophy[42]. For purposes of 'the whole truth', so to speak, I am afraid it will not be enoughto prove them wrong in those terms. My conclusion regarding what is the best view contrary to foeticide (and, by extension, my ultimate conclusions regarding what we have most reason to do in this respect) is to be considered conditional. It is tenable only in

---

to his Creator: in man there shines forth a reflection of God himself", see *Evangelium vitae* (1995), par. 34, italics mine. For a selection of the Catholic Church's doctrine on the especial moral status of human life see par. 355-358, 1700, 1703, 2258 and 2415-2418 of the *Catechism of the Catholic Church* (1993); on its ethical implications for murder see, e.g., par. 2261 and 2268; on abortion, see par. 2270, 2273 and 2274; on euthanasia and suicide, see par. 2277 and 2280. That the wrongness of abortion is on a par with that of other attacks against human life is strongly implied by the following passage: "Whatever is opposed to life itself, such as any type of murder, genocide, abortion, euthanasia, or wilful self-destruction, […] all these things and others like them are infamies indeed. They poison human society, and they do more harm to those who practise them than to those who suffer from the injury. Moreover, they are a supreme dishonour to the Creator" (see *Evangelium vitae*, par. 3, quoting from *Gaudium et spes* (1965), par. 27). Of course, according to the Church's doctrine, none of this applies to what we owe to other conscious beings (*Catechism*, par. 2415-2418): it is justified to treat them in ways it would be very seriously wrong to treat human beings (such as making food and clothes out of them), and whenever some treatment is deemed unjustified (inflicting needless suffering or death, par. 2418) or due ("kindness", par. 2416) it is never for *their sake*, but out of respect for God, as well as for human dignity and quality of life. Indeed, other conscious beings are "by nature destined to the common good of past, present and future humanity" (par. 2415).

[41] Cf. this passage from DeGrazia (2005, p. 280): "For there are no religious theses that are so compelling that it would be unreasonable to deny their truth. […] it is not the case that each of us can be expected to embrace them on the basis of publicly available evidence. […] But in order to be compelling, a moral argument must flow from assumptions that it would be unreasonable to deny. Thus, any persuasive argument against abortion must be persuasive in nonreligious terms".

[42] Such as Finnis (1973; 2011a; 2011b; 2011c; 2011d; 2011e; 2011f; 2011g), Kaczor (2011), Keown (2012) or Lee (2010). Later on I shall examine their views in some detail.

so far there is not a better argument against foeticide based on *true* religious beliefs.

In the previous chapter I sometimes used the opinions of those who accept this account of the wrongness of killing to illustrate the distinctions therein presented. There I suggested a couple of times a view contrary to foeticide which holds that our reasons not to kill a foetus are both agent-neutral and impersonal in character. Truth be told, this is just one possible rendering of this view. Others are defended, as we shall see. This plurality of versions is made possible because of the fact that what is distinctive of Anthropocentrism is not a claim about the *kind* of reasons we have against killing, but rather a claim of *what* it is that gives us such reasons and a claim about their *strength*.

A rather uncouth way to put it would be,

> 'Human life is sacred, so it is almost always wrong to kill people.'

Let us try to expurgate this claim from imprecisions. In the first place we must attempt to discriminate between the living things that count as 'human' for the purposes of this view and those that do not. It is important that this point is clear in order to avoid shallow objections grounded on mere equivocation. For example, regarding those beings which do not qualify as human life: our cells are, certainly, both alive and, in some sense, human, but surely it would be a misrepresentation of Anthropocentrism to take it to attribute special value to, say, a human cancer cell culture in a Petri dish, so that destroying it would be as seriously wrong as killing someone like you or me[43]. Interpreting the phrases 'human life' or 'humankind' in this broad sense would also entail that we should usually refrain from destroying human gametes (i.e., reproductive cells), which are living human cells. Since contraception results in the destruction of egg and sperm, we would have as strong a reason to refrain from killing people as to avoid contraception. Though some who oppose foeticide also stand against

[43] I borrow this point from Marquis (1989, p. 185; 2007b, p. 192).

contraception, I think few would be ready to accept that we have a strong a reason against it as against foeticide. Furthermore, it is not only through contraception that gametes are destroyed, but also through sexual abstinence – each time a woman fails to have sexual intercourse during her fertile periods an egg dies. Yet it seems to me uncharitable to suppose that those who attribute this especial importance to human life are implying that in the death of a gamete there is a disvalue akin, as to the practical reasons it gives us, to that present in the death of someone like you or me. Thus, supporters of this view must mean something else when they speak of 'human life'. We had better construe their claim as stating that the life of those things we name 'human beings', or of which it is sensible to claim that belong to *our kind*, is always valuable, and very much so. Hence, a human gamete is an instance of human life as opposed, say, to a bovine gamete, but it cannot count as a member of our kind. They are, rather, the reproductive cells out of which new human beings are created.

Regarding now the things that *do* count as 'human', one may wonder whether the foetus is actually covered by that concept. As they are ordinarily used, though, the expressions 'human' or 'human being' are deceptively ambiguous. Sometimes by these expressions we mean

> *Human animal*, that is, an organism of the *Homo sapiens* species.

Whereas at other times we mean

> *Person*, that is, a rational, self-conscious being[44].

What Anthropocentrism is intent on proving is that being a *human animal* is normatively significant, for that is what will vindicate its thesis regarding the importance of species membership.

[44] One may find the distinction between persons and human animals, for instance, in Singer (2011, pp. 73-75).

In this respect, it is not hard to accept that the foetus is a human animal, even if in APPENDIX A I problematise this claim regarding the early stages of its development. Certainly, beings like you, reader, or me, are[45]. We are uncontroversial members of that class. Thus, it would turn out that both we *and* the foetus, early in its existence, belong to the same kind. If Anthropocentrism is true, that would show that our reasons against foeticide are given by the very same fact that gives us reasons against killing people. And that is, of course, what writers of the anthropocentric persuasion want to argue for.

Yet we also require a qualification regarding what counts as the 'life' of a human animal. This is necessary because, unfortunately, the term 'life' and its derivatives are also ambiguous. There is a *biographical* sense of 'life' which is the one present in claims regarding whether a life is *worth living* or *worth ending*, that is, whether the aggregate of goods and bads in it turns out to be positive or negative. Only conscious beings, i.e., beings with a mind, can have a 'life' in this sense. But there is also a *biological* sense of 'life', life as a phenomenon evinced by attributes only organisms can possess such as the capacities of adapting to the environment, reacting to stimuli, sustaining their existence, reproducing, etc.[46] This is the sense of 'life' in the anthropocentric claim that human life is always valuable: it matters for good, and it always does, that *there exists* a living member of our species. Affirming that human lives are always worth living, which is a judgement about the personal value a life has for the human organism whose life it is, is certainly compatible with of Anthropocentrism. It might even serve to *ground* it, according to some versions of this view, but it is not what is *meant* by it.

[45] That we are human organisms is certain. What is controversial, as we shall see, is whether we belong to that kind *essentially* or not, that is, whether we could continue to exist without being human organisms. Claiming that we could not is not necessary for Anthropocentrism, but it allows them to say not only that all things that are human are valuable in the way they want them to be, but also that they must be so for the *whole duration* of their existence.

[46] We shall need a more sophisticated account of what distinguishes organisms from other objects soon enough (see CH. III, **§2** and APPENDIX A). For now, I believe this rough characterization will suffice.

What is meant, as per the reasons that have just been given, can be expressed thusly,

Anthropocentrism:

(1) The mere fact that a living organism is a human animal is *always* valuable.

(2) It is valuable in that, except in extraordinary circumstances, it gives us *decisive* reasons not to kill that organism.

On a value-based conception of practical reasons, such as the one I am assuming for the purposes of this dissertation, any plausible account of the wrongness of killing must be grounded on a plausible account of the badness –i.e., disvalue– of death. That is because on any such conception of reasons one of the primary tasks of the moral philosopher consists in identifying the value of the several states of affairs the agent might bring about. Since killing is just causing the death of an organism, in this case this task amounts to identifying what is so bad, if at all, about dying. Anthropocentrism has a ready answer for that –a living organism with a valuable attribute ceases to exist. But *why* and *how* is it as valuable as this view claims? Such as Clara did with her father, it is now our turn to interrogate Anthropocentrism.

§3. *On the impersonal version of Anthropocentrism.*

Let us do it. Why is it usually so wrong to kill beings like you or me? One answer would be,

(3) Being a human animal is *impersonally* valuable.

This is very convenient, since by identifying an organism as an animal of the human kind we know that we have very strong reasons not to kill it without the need to engage in complicated moral calculi regarding the value of the contents of its life. Also, this impersonal version of Anthropocentrism implies that, caeteris paribus, the state of affairs of there being more human animals is preferable to that of there being less. That would help explain why some of those who believe that we always have strong reasons not to kill a human being also believe that we have always strong reasons, not necessarily decisive, to *produce* them –the value of human animals being impersonal in character it gives us reasons both not to destroy existing members of that kind and to create new ones. Anyhow, in so far as the foetus is a human animal, we would have an explanation of why we have such strong reasons against killing it.

At the beginning of this chapter, I drew attention to the fact that the triumph of Anthropocentrism had been bad for non-human animals because this account of our moral status focuses on species membership rather than on individual features, such as cognitive capacities. Anthropocentrism must justify why membership in our species matters more than membership in others. Barring religious considerations[47], how can we account for this attachment of value? I

[47] Perry (2007, pp. 7-13) tries to prove that we all have reasons to accept a Christian foundation (our special relation with a divinity) for the impersonal value of human animals (what he calls 'human dignity'). His argument is that Christian metaphysics is preferable to its alternatives because it is the only one that explains why we have reasons to be moral. Given how the divinity has created the Universe and us it turns out that it is *always* in our interest to act on other-regarding reasons. I do not think this works. In the first place, it assumes that when we ask the question 'Why be moral?' we are asking for self-interested reasons (reasons given by what is good to those we are partial to) to act on moral reasons (reasons given by what is good for estrangers or is impersonally good). Yet this begs the question and he gives no reasons why we should accept that a proper answer should have that form. Secondly, it might be that an appeal to firm normative convictions gives us reasons to choose among different and similarly plausible solutions to complex metaphysical problems. However, Perry is here arguing for a religious foundation to a controversial normative position (Anthropocentrism) by appealing to widely contested theological claims. Unless one is already strongly persuaded by any of this, his argument lacks the firm anchorage in the reader's intuitions which it needs to be plausible.

shall consider two alternatives. We can claim that the special value of human animals is a *basic* principle of practical reason, or practical principle. Or we can claim that it is part of a *larger view*, for example, one which would also account for the impersonal value of Clara's tree or Peter's works of art[48].

Let us start with the first strategy. We could define a practical principle in the following way:

> A proposition expresses a practical principle just in case it is the proposition that we have reasons to try and make some state of affairs obtain.

In this way, we express practical principles when we claim that we ought to brush our teeth daily or that we have reasons to prevent people from suffering and also that there are usually decisive reasons not to kill human animals. Given a value-based conception of practical reason, though, we express such principles too when we claim that it is good for us to brush our teeth daily, or that suffering is bad and also that being a human animal is impersonally valuable. As these examples show, not all practical principles are of the same sort. Some are *basic* and others are *derived*:

> *Basic Practical Principle:* a practical principle is basic just in case it is true in all possible worlds.

Whereas

> *Derived Practical Principle:* a practical principle is derived just in case it is true in some, but not all, possible worlds.

[48] I borrow these distinctions from Tooley (1983, p. 14 and 63; 2009, pp. 24-25), even though I have reworded his definitions as follows.

Consider the claim that it is good for us to brush our teeth daily. This is a fine example of a derived practical principle, in this case a prudential principle. The reason why these are aptly named 'derived' is because they are the result of a valid inference from another practical principle and a contingently true non-practical proposition. Hence Peter can show how it is good for Clara to brush her teeth daily by pointing out that doing so helps prevent all sorts of diseases and that diseases are bad for her. Yet though it is true that brushing her teeth contributes to Clara's health, it could have been otherwise. In some possible world where Clara's counterpart metabolises leftover food crumbs to grow stronger teeth, the principle that diseases are bad for her would give Clara no reasons to brush them.

But let us examine now the claim that suffering is bad. Certainly it seems to be a better candidate for a basic principle. Is there some possible world where suffering is good? I can conceive of a world in which everyone is a masochist and enjoys being in pain. But this is not a situation in which suffering is good, but in which pain causes enjoyment rather than suffering. I can also conceive of many situations in which we have decisive or sufficient reasons to make someone suffer –such as sticking a red hot pin through my fingertip in order to achieve some greatly worthwhile goal. Yet even then I have a reason, even if an ultimately defeated reason, not to stick the red hot pin through my fingertip, namely, that it shall make me suffer. Thus, that we have reasons to prevent people from suffering appears to me as a basic practical principle.

What Anthropocentrism may maintain is that the impersonal (positive) value of human animals is basic in this sense. It would be something indeed to prove that it must be true in all possible worlds. It seems, though, that it cannot be done. First, endorsing this impersonal version of Anthropocentrism has implications many would find counterintuitive. This suggests that there are some instances of human animals which are not as valuable as this theory claims[49].

[49] I borrow the following examples from Tooley (1983, pp. 64-66; 2009, pp. 24-28), whence the interested reader may draw others.

Consider the case of the *philosophical zombie*. It would be a human animal and it would behave exactly like a human animal –it would laugh if tickled, cry if hurt, express its delight in a sunny day and it even might appear to seek revenge if wronged. However, it would have no mental states, so that, in spite of its behaviour, it would feel neither joy nor pain nor pleasure nor hatred. Impersonal Anthropocentrism would that this organism is as valuable as those other human animals who are actual subjects of experiences, and for the same reasons. I find it hard to believe that I would have the same strong reasons not to kill a philosophical zombie as not to kill an ordinary human animal. Or that I ought to find the existence of the former, caeteris paribus, as desirable as that of the latter. Presumably that is because being a conscious entity is very normatively significant.

Some may object that the notion of a philosophical zombie is incoherent, so it might be good to disprove that claim (3) is a necessary truth via other counterexamples. Human animals need not be philosophical zombies for them to be both alive and not conscious beings. It is enough that they lack those areas of the brain responsible for conscious experience. Consider, for instance, anencephalic infants. These are babies born without cerebral hemispheres, and thus, presumably without the capacity for conscious experience[50]. Consider also

[50] See Steinbock (2011, pp. 23-31) and the literature referenced in APPENDIX B for a discussion of whether a functional cortex is necessary for consciousness and whether anencephalic infants can or cannot be described as conscious beings. Another interesting case of abnormal human development is that of the *foetus in foetu*. Sometimes twinning goes wrong. One of the twins continues to develop in a normal way, except for the fact that it envelops its brother with its growing body. The enveloped twin, the *foetus in foetu* (the foetus *within* the foetus), eventually stops developing. Yet it is alive, inside a placenta, and connected to its brother's organism by an umbilical cord. The other twin may become a normal adult human being –normal, that is, save for carrying its underdeveloped, living brother inside (Meyers, 2010, pp. 83-100). Some deny the *foetus in foetu* originated as a twin of the bearer, and claim that this is enough to exclude this case as a counterexample to impersonal Anthropocentrism. Supposing that they were right in their factual claims, this entity would not have the same impersonal value as a human animal. However, it doesn't follow from this that it cannot be used as a counterexample. It can still be used even if it is merely an *imaginary case*.

the case of people in permanent vegetative state (PVS), and which can be kept artificially alive. It is undeniable that they are human animals. Yet again, it seems that we do not have so strong a reason to desire their existence as that of those humans who are subjects of experiences, and that the usually decisive reasons not to kill the latter are absent in their case.

Impersonal Anthropocentrism faces these counterexamples because it disregards the importance of the personal value of one's life for the wrongness of killing. That is, it tells us that we have strong reasons not to kill human animals for which nothing can be good or bad, like PVS patients or anencephalic infants. Yet it gets even worse, for it tells us that we may have decisive reasons not to kill human animals whose lives are *not* worth living. Imagine someone the rest of whose life will be utterly miserable, perhaps because she is in constant agony. Suppose, moreover, that this person's pain is so intense that she is unable to enjoy anything which otherwise would have constituted a good for her. Also, let us assume, there is nothing that we can do to remedy her situation. In these circumstances most of us would claim that such person has at least sufficient reasons to commit suicide and that we all have sufficient reasons to comply with her petition to assist her in case that she asked. Anthropocentrism need not deny that her misery gives her reasons to kill herself, but anyhow on this version it must claim that she would still have very strong reasons to keep herself alive, which include reasons for us not to assist her.

At this point I believe Anthropocentrism has two options. One would be to claim that we always have decisive reasons against suicide and voluntary euthanasia, which I find deeply implausible. Suppose that because of the foreseeable distribution of goods and bads in my future, it is certain that beyond some point in time my life will be worth ending, from a self-interested perspective. This might occur, for example, due to a medical condition which inflicts excruciating pain for most the time months before it causes death. Since beyond that point in time the overall value of my life would be negative, I would have strong

prudential reasons to end it. Assuming that no further bads would accrue because of my death, such reasons would be decisive. But, as I said, this is incompatible with one interpretation of what Anthropocentrism tells us we have most reason to do in these cases. Yet, as I have defined this position, it merely claims that our reasons not kill human beings are *usually* decisive. Also, any plausible theory must say so in order to account, for example, for justified self-defence. So Anthropocentrism could alternatively claim that maybe the reasons given by the fact that someone's life is not worth living are stronger than the reasons given by the fact that she is a human animal. Thus, a sufficiently large aggregate of bads in one life can outweigh its impersonal value and give us, say, sufficient reasons to commit suicide. But if Anthropocentrism goes this way, one may legitimately wonder why the same cannot be true of an aggregation of goods and whether, in fact, it is the personal value of a life what gives us our main reasons not to end it.

The force of the impersonal value that this version of Anthropocentrism attaches to human animals doesn't stop in giving us reasons against killing members of our kind, but, as I suggested earlier, it also gives us equally strong reasons not to prevent their existence. On a simple, merely additive, account of the aggregation of value, the strength of our reasons not to kill, or to produce, an extra human animal will remain constant and will not depend on the amount of human animals in existence. On a more complex account of the aggregation of value it may turn out that after we pass a certain quantity-threshold of humans, the existence of an extra human animal adds increasingly less value to the outcome. On any account of aggregation, though, the following is true: our reasons not to kill a human animal in a world where n exist are as strong as our reasons to produce a human animal in a world where n–1 exist. And if a *caeteris paribus* clause is added, whenever our reasons not to kill in the first scenario are decisive, our reasons to produce in the second scenario *also* are. This actually means that our reasons not to engage in actions that prevent the existence of future persons, like celibacy, sexual abstinence and the use of

contraception are as strong as our reasons not to kill people. I think all this is false.

I have been assuming that authors persuaded by this version of Anthropocentrism would not bite the bullet. But they can accept implications such as that we have usually decisive reasons not to end our own lives, even when they are not worth living and even when no other bad consequences would ensue. Thus, what many find implausible may, nevertheless, be perfectly acceptable for supporters of Anthropocentrism. So let us assume, for the sake of the argument, that claim (3) expresses a *valid* practical principle. One way to show that, in any case, it must be derived would be by producing similar such principles that also strike us as intuitively valid and see whether they can all be inferred from a more general proposition.

Imagine that some extraterrestrial being came to pay us a visit at long last[51]. It is not a human animal, for its genome is unlike ours and it cannot interbreed with us. Nevertheless, it is apparent by the technology it employs that it has cognitive capacities at least on a par with ours. Also, it learns to communicate with us, and we discover it is capable of quite solid moral and aesthetic judgements. Presumably, even if we are persuaded by Anthropocentrism, we would want to claim that it merits a fundamental treatment similar to the one a human animal deserves. Yet, if we do, it would be rather unparsimonious to claim that there are two basic moral principles here regarding, respectively, the impersonal value of human animals and extraterrestrials. And especially when we have come to accept the need to develop our new principle by recognizing in our acquaintance with the extraterrestrial the salience of attributes such as its cognitive capacities and its moral and aesthetic faculties, which human animals also have. It is more plausible, then, to suppose that both principles derive from a third one attaching value to the possession of said faculties or capacities. This

[51] Here I merely reproduce an argument common among critics of Anthropocentrism. Instances of it can be found, for example, in Boonin (2003, pp. 23-25), Tooley (1983, pp. 67-68) or Warren (1973).

basic principle, however, could still cover all the cases Anthropocentrism would like it to cover –like foeticide–, since it could appeal to potential capacities, or the capacities stereotypical members of the species actually possess[52].

Another case can be made for the claim that the principle that human animals are impersonally valuable must be derived. It consists in showing how the attribute it appeals to, membership in a species, is unapt to possess, in and of itself, rational significance[53]. Membership in a species, as membership in any class, is determined by a set of individually necessary and jointly sufficient conditions. Thus, if members of a species are valuable *qua* such it must be because some or all of the attributes these conditions refer to are rationally significant. Since the classification of living organisms in species is part of the task of biology, it is to experts in that field that we ought to ask what criteria they use for distinguishing among species. I shall consider here two such criteria, namely, the capacity for interbreeding and a common genotype. First, suppose that organisms belong to the same species if they can interbreed. The problem is that sometimes this criterion leads to contradiction. Suppose that animals of set A can interbreed with animals of set B, animals of set B with those of set C, and these with animals of set D. As per the interbreeding criterion, animals of set A belong to the same species as those of B, animals in B are members of the same species as animals of set C, and animals in this latter set are of the same species as those in set D. Assume also that the polyadic attribute 'being member of the same species as' is transitive. We could then infer that animals of set A are of the same species as those of set D. And yet, there are in fact

[52] Sometimes the issue is presented as if what Anthropocentrism ought to prove is that being a human animal is a *necessary* condition for having the special value we usually recognise each other. If that were the case, the argument just developed could prove this view wrong, since it forced us to admit that members of some other species were equally valuable. But as Finnis (2011d, p. 224), Keown (2012, p. xx) or Lee (2010, p. 60) acknowledge, Anthropocentrism need only claim that membership in our species is *sufficient*. It can happily admit that membership in other ontological kinds is also sufficient, and on the same grounds, for having such especial value.

[53] In subsequent paragraphs I shall be following McMahan (2002, pp. 212-214).

instances where that is not true. This may happen because though animals of set A live in a territory adjacent to those of set B, these in one adjacent to the territory where animals of C live and so forth, animals in A and D are geographically isolated. But it may also happen because animals in A and D are *temporally* isolated. Each of us belongs to the species of our biological parents, they to that of our grandparents, and so on. However, if we travelled back in time we could not interbreed with our amphibian ancestors. So, do animals of A and of D belong to the same species? The interbreeding criterion, unfortunately, forces us to answer both affirmatively and negatively. Besides, the capacity for interbreeding does not seem normatively significant, or at least not enough so as to ground the very high value impersonal Anthropocentrism attaches to human animals.

Perhaps possession of a common genotype fares better as a criterion for distinguishing among species. Again, though, we must ask ourselves how possession of this attribute could matter in the way this view wants it to. For one, sometimes there is very little difference between the genotype of different species which are phenotypically very dissimilar, such as happens with humans and chimpanzees[54]. Also, we know it is possible to combine genes from different genotypes belonging to different species, thereby creating transgenic animals, for example transgenic chimpanzees with human genes. We could get a chimpanzee zygote and substitute just one of the genes in each cell with a human one. Alternatively, we could extract its whole DNA except for one gene and substitute it with a complete human genotype minus one gene. But those would be only the two extreme cases of a spectrum of possibilities in each of whose scenarios the chimpanzee would have one extra human gene than the animal in the preceding scenario. A supporter of Anthropocentrism would need to determine how much human genotype is enough in an organism for it to be valuable, and that requires identifying the earliest point in the spectrum where the animal can be considered human. If we asked for a justification for her choice, she could only point out to differences in the quantity of genes.

[54] See Hull (1998, p. 389).

However, that does not seem as important as, for example, what human genes are present. We can conceive of an animal with more human genes than another, and yet with the latter possessing the sort of genes that direct the development of high cognitive capacities and the former possessing those that direct the development of a distinctively human appearance. It does not seem plausible to say that the degree in which a genotype is human has any significance –it gives us no reasons to prefer the animal with more human genes over the other.

I believe we must reject the idea that claim (3) expresses a basic practical principle. Even if we are unconvinced by the counterexamples here developed one must accept that it is derived –membership into a species cannot be necessarily valuable, so we need show how being a human animal relates with some further valuable property. This is the second strategy I mentioned earlier in this section. In CHAPTER I, when explaining the notion of impersonal value, I tried to exemplify it by suggesting the possibility that it is a sort of value works of art and works of nature possess. If we can explain the impersonal value of human animals as an instance of the value natural entities in general have, and the reasons we have not to destroy them, perhaps Anthropocentrism can defend itself from the charge of giving arbitrary importance to membership into our species[55].

One of the best known attempts to reconstruct the derivation of claim (3) from more general practical principles has been Ronald Dworkin's[56]. He proposes a

> *Genetic Account of Impersonal Value:* there is a set of processes such that
>
> (a) Their products are impersonally valuable, and

[55] For the charge that Anthropocentrism is speciesist see McMahan (2002, p. 214), Singer (2011, pp. 48-53 and 94-122). This problem is also briefly discussed in Glover (1977, p. 50).

[56] Dworkin (1994, pp. 81-84).

(b) They are valuable *merely* because they satisfy the predicate 'embodying an important creative process' (and not because of any other attribute they possess).

We are also told that two such classes of processes are natural and human creation, and that some instances of the latter and at least many instances of the former have as a result entities of impersonal value. Supposedly, this account accommodates our intuitions about how it is proper to treat works of art, which constitute paradigmatic products of human creation. Dworkin then notes how we use a similar language to explain how we should respect the results of natural processes of creation. Admittedly, since nature's creative processes are non-intentional, the comparison with man-made works of art is imperfect. However, it is allegedly the best way we have found to convey a deep intuition that the *investment* that through aeons nature has made in its creations –for instance, animal species– is worthy of our respect. Finally, it is remarked how human beings are, in this sense, nature's creative masterpiece, produced by means of a long process of natural evolution. Thus, "the life of a single human organism commands respect and protection […], no matter in what form or shape, because of the complex creative investment it represents and because of our wonder at the divine or evolutionary processes that produce new lives from old ones […]"[57].

As a reason to account in this way for the impersonal value of works of art or of species Dworkin claims that grounding it in other attributes different that their relation with their process of creation would have unsavoury implications. We would have to admit that the value of artistic and natural creations is *incremental* –that is, that we have reasons to prefer the outcome in which there are more of them to the alternatives, all else being the same. However, he does not believe that we value works of art or animal species in this incremental way. Rather, the impersonal value of these things is of the sort he calls *sacred* or *inviolable*. Regarding works of art, we deem it a tragic loss when one of them is

[57] Dworkin (1994, p. 84).

destroyed, but, according to Dworkin, we do not have a corresponding similar attitude, except perhaps with masterpieces, in case we fail to create them. As to species, we do not *always* believe that it is better when there are more human beings[58]. So, he argues, all these are objects we merely have reasons not to destroy, but not reasons to create. Purportedly, by claiming that these things are impersonally valuable "because –and therefore only once– [they] exist"[59], that is, because of how they have been created, he can explain why we have one sort of impersonal reasons and not the other.

I do not think this works. According to Dworkin animal species or works of art are valuable because they embody or represent the investment human beings or nature has put in their creation. Imagine now a paradigm instance of work of art in which gigantic effort has been invested, such as Michelangelo's frescoes in the Sistine Chapel. Let us reprise our role as World Maker. Now we must choose between two possible worlds to create. In the first there will be someone who will paint the frescoes of the Sistine. In the second one, such frescoes will not exist. I believe we have reasons to prefer the first world and, caeteris paribus, decisive reasons. If that is true, it shows that what Dworkin's account implies is not that we do not have reasons to desire more works of art, but that we have reasons to desire that there be successful *artists*, that is, people who put effort in producing such works and manage to create them.

Yet his account also implies the following. Suppose our World Maker is presented with a different choice. One alternative is to create a world in which Michelangelo's frescoes of the Sistine do not exist. The other is to create a world in which such frescoes exist, but are not the product of an artist but of the random rearrangement of particles of matter. Also, in every other respect these worlds are exactly alike. According to Dworkin's account we would have no

[58] Here Dworkin (1994, p. 70) seems to be worried about the possibility that his account implies that we have reasons not to prevent the existence of human beings which are as strong as our reasons not to kill them.

[59] Dworkin (1994, pp. 73-74).

reason to prefer the second over the first. Again, I believe this is implausible. I believe we have decisive reasons to prefer the second of those worlds. I do not mean to deny, or to assert, that the artist's effort is part of what constitutes the value of a work of art. Yet it cannot be all that there is to it –some further attribute of the piece must have significance. For example, we could say that whenever something is aesthetically valuable it contributes to the impersonal value of the world it populates whenever it improves the lives of someone who experiences it.

The analogy with the impersonal value of works of art, then, may not be the best argument for *natural investment* as the grounds for the impersonal value of species. Such analogy requires, in order to succeed, a sound case for the significance of human investment for the value of art. We just saw the deficiencies of Dworkin's strategy in this respect. It also requires, though, showing that natural creative processes are sufficiently similar to human ones. But even if artistic processes were a source of impersonal value, they are so unlike natural creative processes that the analogy is hardly sustainable. Most saliently, human creative processes can be intentional, whereas natural processes cannot. It remains mysterious, then, how evolutionary processes are the source of a supposed impersonal value of biological species instead of that source being, say, the common attributes of their members.

For the sake of the argument, though, let us assume that Dworkin's proposed reconstruction of our intuitions is a sound explanation of the impersonal value of some natural phenomena. Are we then justified in claiming that the impersonal value of members of our species is especially superior to that of members of other species and in the way Anthropocentrism requires? There is an objection that this *Genetic Account* rests on a misunderstanding of what the process of natural evolution entails[60]. The claim that we are the supreme product or "masterpiece"[61] of that process is unjustified, for it is false that natural forces

[60] As suggested by Boonin (2003, p. 31).

[61] Dworkin (1994, p. 82).

were somehow ordered to culminate in us. Members of all species, including us, are a result of a natural selection of randomly varying individual traits. And all members of existing species, not only ourselves, are in the same way the end of an aeons-long process of evolution, which they managed to survive. Thus, there is nothing in the investment account of our impersonal value that grounds a special consideration to human animals as opposed to non-human creatures.

I know of no other attempts at justifying the claim that human animals are impersonally valuable as a derived practical principle, save, of course, religious ones. If the arguments I have developed here are sound, Dworkin's proposal also fails to do that job. Additionally, we saw how implausible it is that it can be considered a basic principle. A claim about values and their reasons that cannot be defended either as basic or as derived must be rejected. We shall have to see whether there is a more cogent version of Anthropocentrism.

§4. *On the personal version of Anthropocentrism.*

Some authors try to rescue the special importance of membership into our kind by dropping the claim that being a human animal is impersonally valuable and stressing how it affects the personal value of the organism's life. This move has, I believe, the advantage typical of those accounts that derive our reasons against killing beings like you or me from the personal value of our lives, namely, that of accommodating the powerful intuition that what primarily matters in these cases is how bad death is for the one who dies. I say 'primarily' because someone who endorses an account such as these can unproblematically admit that we have additional reasons against killing and that these may be impersonal in character.

As we shall see, many authors who favour personal accounts of the badness of death believe that in order to assess how much the deceased has lost in dying we must judge how good or bad the life she would have otherwise lived would

have been for her. Of course, how this is to be done is very contentious, but one would like to say that some conscious beings have had lives more worth living than others. One reason for this is that different individuals are not equally fit for the various goods and bads that constitute the personal value of a life. That is because psychological capacities establish a set of constrains to how personally valuable a life can be by fixing the kind of goods and bads that can be present in it. Thus, my dog cannot obtain the good of choosing and carrying out a life plan, whereas I can. Conversely, he cannot suffer the bad of failing to achieve life-long goals because of social injustice, ineptitude or misfortune. And many times I think I shall never experience something like his moments of absolute contentment, for so much more is needed to satisfy me. But, on the accounts I have now in mind, even the lives of individuals with similar psychological capacities are not always equally worth living, because fortune varies across individuals. At one end of the spectrum of fortune some will lead very flourishing lives, whereas at the other end we may find people whose life is rather worth *ending*.

When one derives her reasons not to kill people from the personal value of their life, and given the fact that some lives are more worth living than others, the unpalatable conclusion may be reached that we have stronger reasons to kill some people than to kill others. For many, this is troublesome, because they believe that the strength of our main reasons not to kill any of us should be the same, even if the strength of the other reasons that enter in the balance may vary. Was it not what having a *right to life* was all about? Different authors have pursued different strategies to remedy this. In this section I shall review some proposed by authors of the anthropocentric persuasion, who see a chance to vindicate the importance of species membership.

What these anthropocentric authors claim is something like,

> (4) The life (biological sense) of a human animal is *always* very personally valuable.

These authors believe, moreover, that we are essentially human animals, that is, that we are essentially a specific kind of living organism[62]. This means that our life is especially valuable from the beginning of our existence until its end, since we necessarily belong to humankind for as long as we persist. On this account, also, the valuable object which gives us reasons not to kill is not the *contents* of someone's existence, its aggregate of goods and bads. It is the *mere existence* of a human animal –its being alive– what is conceived as good. This, I believe, is what distinguishes these personal versions of Anthropocentrism from other personal accounts of the wrongness of killing and the grounds on which they can justify an equal right to life for all human animals[63].

One variant of personal Anthropocentrism is that endorsed by Christopher Kaczor[64]. Imagine that it could be shown that membership into our species determines the personal value of the life of each of its members. Since this is presumably an all-or-nothing property, either a being belongs to it, and then its life has the special value of a human animal's, or it does not. Moreover, if Kaczor is right, and assuming the foetus is a human animal, we will usually have decisive reasons not to kill it. Succinctly put, his argument is that our species, just like any other, possesses a nature shared by all its members. That nature is what determines what the good, or flourishing, of beings of that species consists in. Our species has a rational nature, which means that all humans are directed to the most valuable goods there are, namely, freedom and rationality[65]. That explains why killing a human animal is worse than causing the death of any other living thing –it was an organism aimed at the

[62] See Finnis (2011a, pp. 280-281; 2011d, p. 219), Kaczor (2011, pp. 105-120), Lee (2010, p. 71).

[63] Finnis (2011d, p. 219). I shall discuss said other accounts in CH. III.

[64] See Kaczor (2011, pp. 91-102).

[65] See Kaczor (2011, p. 95).

relatively highest form of flourishing and it has been bereaved of something, its life, necessary for it[66].

Now, this account of the importance of species membership has several advantages over the impersonal accounts reviewed in the previous section. It does not claim that it is a basic practical principle that species matters. Also, it is based on how belonging to our species affects how good our life can be for us, which seems more promising than appeals to the creative processes of nature. Certainly, though, something else is needed to render credibility to a view such as Kaczor's. Since he appeals to such things as species and their nature he needs to tell us more about what they are and how they can possibly be important.

Let us start with his conception of *species*. I would argue that it is inconsistent. There are a couple of revealing passages where it is clear that he does not adhere to a genetic criterion of species membership[67]. Instead he adheres to two –a *genealogical* criterion and a *capacities-based* criterion. According to the first criterion one belongs to the same species as her parents. However, according to the other species are distinguished by their powers or capacities. We can observe how these criteria can clash by considering one of Kaczor's own examples. Imagine a cow with human-like cognitive abilities, perhaps due to genetic manipulation of the zygote from which it developed. According to the genealogical criterion, it should be considered an ordinary cow. Yet, as Kaczor himself admits, as per its capacities it ought to be considered a new animal species, with a distinct nature. In this and other similar cases, he solves the contradiction in favour of the capacities-based criterion.

[66] Kaczor (2011, p. 97) also claims that life functions as a non-instrumental good, in that it is part of what having a good life consists in for an organism. However, it seems to me that it is the instrumental value of life what varies from species to species, since such kind of value tracks the worth of the goods in which the flourishing of each species consists.

[67] See Kaczor (2011, pp. 21-23) for this and the following paragraph.

This makes some sense. One may think that the capacities-based criterion has *priority* over genealogy because it is what Kaczor believes membership into a species actually *consists in*, whereas identification of the parents' species is merely a good proxy –the answer it gives us correlates with the answer given by the true criterion in most cases, but not all. For example, we have seen how at some point it ceases to be true that we belong to the same species as our ancestors. But if organisms belonged to the same species as their parents, assuming all living things have a common ancestor, they would all be members of the same species. Kaczor cannot admit that, because then his species-specific account of an organism's flourishing would collapse. Since he rejects a genetic criterion of species membership, appealing to a capacities-based criterion is the way to escape from that implication. Notwithstanding that, in some other passages Kaczor gives priority to the *genealogical* criterion. Indeed, he insists that even if a being has not the requisite psychological capacities for rationality it must be considered as a member of our species insofar as its parents were human animals.

This shows, I believe, the arbitrary fashion in which Kaczor distinguishes among species. Since he wants to defend the claim that

(a) Whatever is good or bad for an individual being is determined by the nature of its species.

whenever he is confronted with high psychological capacities in a non-human animal he must appeal to a capacities-based criterion and claim that a change in species has obtained. If he did not, he would have to admit what are for him counterintuitive implications as to how we are morally permitted to treat an animal which has proven itself rational. But since he also wants to defend that

(b) All human animals possess a rational and free nature.

when confronted with counterexamples about the absence of high psychological capacities in a human animal, he must retreat to a genealogical criterion. By doing so he can still claim that these beings belong to the human species and partake of a rational nature, which is of course, what he wishes to conclude.

But the problem runs deeper, since the cogency of Kaczor's claim that the "nature of the species'" determines the personal value of a human animal's life is undermined by his use of the genealogical criterion to defend claim (b). To see why let us say something more about how to identify the nature of a species. According to Kaczor we can do that by observing how "healthy, mature members of the species function"[68]. For example, paradigmatic members of our species are free and rational, so that means that our species possesses a rational nature. Let us accept his use of the genealogical criterion to include cognitively impaired children of human animals as members of our species, which is what the genetic criterion Kaczor rejects would anyhow demand. One may legitimately wonder how the psychological capacities of mature, healthy human animals can give us reasons regarding how to treat humans that lack such capacities as freedom and rationality.

Kaczor may argue that there is a sense in which all human animals are *actually* rational beings, even if not all of them can exercise rationality. This is the sense that even in those cases where normal human flourishing is "impeded by unfortunate circumstances, such as disease, or by deliberate choice of others" it still consists in achieving the goods of freedom and rationality[69]. But consider the anencephalic infant, or someone who has irreversibly lost her capacity for consciousness. In what sense is it aimed towards rationality when it lacks the proper organs for it? Kaczor suggests that this is the only way to account for our intuition that cognitively impaired human animals are unfortunate in a way in which other animals with comparable psychological capacities are not –they

[68] Kaczor (2011, p. 23).

[69] Kaczor (2011, p. 95).

cannot “fully flourish as the kinds of being they are”[70]. Yet it would be to no avail to Kaczor if we accepted that the flourishing of anencephalic infants, and all human animals, consists in achieving freedom and rationality. I cannot see how it can be true that in killing an anencephalic infant I am causing it to *lose* the good of rationality, or any of the goods which ordinary human beings pursue, no matter how oriented towards it they were. That is because although we kept the infant alive, it would *never* experience those goods, not even desire them. Even if species determines the right class of *other lives* for comparison in order to assess how relatively fortunate an individual is, in order to assess how bad death is for that same individual we must check how *her own* life would have been if she had continued to exist[71].

Kaczor seems to be aware of this objection. Why, then, not abandoning a species-specific account of flourishing in favour of one that does a better job of tracking what each individual actually loses in dying? He does suggest that we cannot be expected to know what lies in store for each of us in the future, and thus to adopt an individual-specific account of flourishing in our everyday moral judgments.[72] Here the species-specific account appears to be useful as a rule of thumb –since we cannot know for certain the personal value of people’s lives, it is argued, we had better assume that all human animal’s have it equally high. This is a bad defence of Kaczor’s position. For one, if he is trying to produce ideal moral theory he cannot defend his position by appealing to what might be a good non-ideal moral principle to use when we are ignorant or uncertain about the relevant facts. Secondly, though it does follow that in our everyday practical judgements we will have to use frequently an approximate criterion, what does not follow is that it must be species-specific. That is because we can do better than that and draw finer distinctions. Granting that there are difficult cases, we

[70] Kaczor (2011, p. 101).

[71] It must be noted that Kaczor’s view also implies that we always have strong reasons against committing suicide, and against assisting others to do so, even though our future lives may be absent of whatever makes them worth living. As with the impersonal version of Anthropocentrism, I must say I find this hard to accept.

[72] See Kaczor (2011, p. 102).

can nevertheless clearly identify some human animals as capable of exercising rationality and some as clearly incapable of doing so. There is no reason why in the latter case Kaczor's account should be employed.

Thus, Kaczor's conception of species membership is inconsistent and he fails to show why what he calls the 'nature' of a species is normatively relevant in the way he wants it to be. Also, there is no reason, I believe, to accept his account of the wrongness of killing even in our everyday judgements as an approximation to the truth. Thus, the claim that

> (4) The life (biological sense) of a human animal is *always* very personally valuable.

cannot be justified in this way.

Other authors however, have followed different strategies. One of them is John Finnis. He believes that what practical reason requires is respect, both in our choices and in our dispositions, to all the basic goods that constitute human flourishing[73]. Life, in my biological sense, is one of these. That is, being alive is not merely *instrumentally* good in that it is necessary for the enjoyment of the rest of basic human goods, but it is also good *for its own sake*. Now, a choice against a basic good, such as life, can be either direct or indirect[74]. One chooses against life *indirectly* when the death of a human animal, though foreseen and accepted, is not intended. According to Finnis, our reasons against indirect killing need not be decisive in all cases. Direct choices against life, for their part, are those in which an agent intends the death of a human animal either as an end or as a means to a further end. The prohibition against direct killings is absolute –we always have decisive reasons against intending the death of a human animal. This prohibition against direct killing would merely be an instance of a wider principle according to which we always have decisive

[73] Finnis (1973, p. 125; 2011g, p. 295).

[74] For the direct/indirect distinction see Finnis (2011g, pp. 297, 299-301).

reasons against choosing “to destroy any basic human good —any intrinsic aspect of *personal* well-being— for the sake of any ulterior good, however important”[75].

In the passage I last quoted, Finnis talks about ‘personal’ well-being, by which he means the well-being of a *person*. Now, Finnis characterises persons as those beings with such higher intellectual capacities as freedom, rationality or deliberation[76]. By *capacity*, however, he means something quite different from the ordinary usage of that term. A being not only possesses a capacity for intellectual activity when it is able to engage in an intellectual act, but also when it has “a genetic constitution normal enough to provide, or develop sufficiently to provide, at least the organic basis of some intellectual act [...] even when too impaired to perform such an act’, for example, by disease, accident or senility”[77]. This second-order capacity to acquire first-order capacities by sheer metabolic development is what Finnis calls a *radical capacity*[78]. According to Finnis, this justifies the claim that all human beings are persons[79], that their lives are much more valuable than those of non-personal beings and that our reasons against killing them are usually decisive in the manner described above. Assuming foetuses are organisms of the specific kind we are, all this is also true of them.

[75] Finnis (2011g, p. 302), italics his.

[76] Finnis (2011b, p. 288; 2011d, pp. 220, 238).

[77] Finnis (2011d, p. 220), but see also Finnis (2011a, p. 284; 2011b, p. 288; 2011d, pp. 224-225, 227-228, 238-240).

[78] Finnis (2011a, p. 284 and endnote †).

[79] He claims that there are no human individuals who are not persons (2011d, pp. 224-225). Yet are not there some who just lack a sufficiently normal genetic constitution? There are products of human conception which Finnis accepts cannot be considered persons, even on his definition, but which cannot be considered *human animals* either. He mentions hydatidaform moles (2011d, p. 220), which is the name given to abnormal embryos characterized by having only paternal genes. They result in empty placentas usually aborted before the pregnancy is even detected (Sadler, 2012, pp. 48, 364).

Yet what reasons do we have to accept Finnis' account of the wrongness of killing? First of all, we might wonder why his account of personhood in terms of radical capacities is preferable to one that characterises persons as those beings with a first-order capacity for rationality. Since this is a capacity that we might have not yet acquired –because of immaturity– or might have lost –by disease or accident or senility–, it is possible that at some time a human animal is a person, whereas at some other time it is not. Consider, again, someone who is in a PVS because she has lost all the areas of the brain on which the capacity for consciousness depends, but who is nevertheless alive. We could say that there lies a human animal who is no longer a person. Now, Finnis objects against this possibility that it implies a form of dualism, a theory claiming that there is one substance (a person) inhabiting another (a body), and that dualism is untenable because it cannot account for our sense of a unitary self. On dualism, there is no single entity to which the various attributes usually ascribed to the self can be predicated. Rather some –like metabolism, motor function, or sensation– are ascribed to the body, whereas others –like developing a complex argument– are ascribed to the person. Notwithstanding the force this objection might have against dualism, I do not believe that in affirming, say, that a human animal is no longer a person we are necessarily committed to the claim that where there were two substances now only one remains[80]. Instead we may conceive of personhood as *nothing more* than the possession by an entity, like a human animal, of certain first-order psychological capacities. Thus, when we claim that someone in a PVS is no longer a person we may be just committed to the claim there are certain attributes –psychological capacities– and a certain being –a human animal– such that the latter possessed the former but it presently does not. Hence, we need not be worried by Finnis attack against dualism if we choose to define personhood in terms of first-order capacities.

Perhaps, however, Finnis can defend his account of personhood by showing that it possesses a normative significance which the proposed alternative lacks.

[80] See Finnis (2011a, p. 285; 2011c, p. 318; 2011d, pp. 220-221, 238).

Indeed, he suggests that if we accept that personhood is not essential to a human animal but merely contingent it cannot have the importance in practical reasoning with which we usually bestow it, what he calls the "significant depth" and "dignity" of one's being a person[81]. The force of this objection, though, is entirely dependent on the persuasiveness of Finnis's conception of the practical importance of personhood. Well, how persuasive is it? Let us consider first its implications. One of them we have already encountered –since we always have decisive reasons against intending the death of a human animal, we always have decisive reasons against suicide or against complying with a request for euthanasia. No matter how unable to participate in other basic human goods (that is, no matter how miserable someone's life might be) a human animal always participates in the basic good of life, attacks against which are absolutely forbidden. Also, always according to Finnis, since basic human goods are incommensurable, it is irrational to judge that the badness in one's life may outweigh the good of being alive and render suicide, all things considered, morally permissible[82]. Another entailment of Finnis's view would be that the life of an anencephalic infant is as good for it as it is for some self-conscious and rational being. He must conclude this since as far as we know, anencephaly is not due to a genetic condition that prevents the development of the organic basis of intellectual activity[83].

What these entailments suggest is that Finnis's conception of the importance of personhood is unpersuasive enough to cast doubts both on his account of the goodness of life and on his account of personhood based on his notion of radical capacity. As to the latter, consider again the case of the anencephalic infant. Assume that we discover that some cases of anencephaly result from a genetic defect. These are embryos which lack the genes necessary for the development of the cerebral hemispheres. Babies with this condition would be

[81] Finnis (2011d, p. 225).

[82] Finnis (2011c, pp. 318-319; 2011d, p. 236).

[83] A sign of this is that most cases can be prevented provided the mother adopts certain habits during pregnancy (Sadler, 2012, pp. 137, 361).

entirely like ordinary anencephalic infants in that they would lack the same organic structures and, consequently, the same psychological capacities. On Finnis's account, babies with the 'new' anencephaly would not be persons whereas babies with the 'ordinary' one would be, with the corresponding difference in the value of their lives. I simply find it implausible that this etiological difference may have such tremendous normative significance.

Consider, though, an objection Finnis may raise. Abandoning the radical capacity account, we obtain the result that there are many children and adult human beings who, because of some cognitive deficiency, cannot be considered persons. These human animals possess psychological capacities similar to those we observe in other sentient species. It is surely unacceptable to claim that we have reasons to treat them in the same way we treat other animals. After all, we consider that, given their capacities and the kind of valuable life these animals can lead, we usually have sufficient reason to kill them for food, clothes or for sport. Does this mean that we can permissibly kill human non-persons for the same reasons? If this objection were to succeed it would deal a terrible blow to positions rival to Finnis's. But it need not succeed. For it rests on a certain assumption about how animals may permissibly be treated which it is not necessary to accept. Perhaps it is animals that ought to be treated in a way more similar to how we currently care for human non-persons[84]. If we recognize that we have been wrong in our dealings with non-human animals, we can argue that this objection rests on a mistaken assumption.

But let us return to the cases of suicide and euthanasia. If we obtain the implausible conclusion that we never have even sufficient reasons for suicide or for complying with a request for euthanasia is because Finnis follows the tradition of considering life an intrinsic basic good and basic human goods incommensurable. For assume life were commensurable with the rest of basic goods –then it would not be irrational to suppose that the fact that the contents

[84] This is the strategy McMahan calls *Convergent Assimilation* (2002, pp. 228-232).

of someone's existence consists juts in bads outweighs the goodness of her life. Or assume moreover that life's goodness were merely instrumental. Then a human life would be just as good as the positive net sum of its contents was. Sadly I lack the space to discuss Finnis' defence of the incommensurability of basic human goods, which is part of a more general objection against what he calls "consequentialist" or "proportionalist" moral theories[85]. What I will discuss, though, is his contention that considering life to be merely instrumentally good leads to unsavoury implications.

So far I have been considering voluntary euthanasia. Finnis expresses his fear that, once we accept that we might have sufficient reasons to commit voluntary euthanasia on the grounds that someone's life is worth ending, those same grounds will lead us inexorably to conclude that in some cases our reasons for other kinds of euthanasia will be sufficient too[86]. Finnis here distinguishes between *non-voluntary* euthanasia and *involuntary* euthanasia. The former consists in killing someone whose life is deemed worth ending even though she has *not* manifested her will to that end, but neither against it. The latter occurs in those cases when the victim had expressed his will *against* being killed. Finnis notices that even assuming that autonomy has value, and that therefore we have reasons to respect someone's autonomous will or not to procure her death if we are ignorant of her desires to that respect, one must accept that in some cases the disvalue of someone's life will outweigh the value of her autonomy and that in those cases our reasons for killing her might be sufficient.

I believe Finnis's reasoning here is more or less correct. It could be sophisticated a little bit more by adding that there might be other reason-giving facts beyond the personal value of someone's life and her autonomous will, such as, for example, the effects of our action on other people. All in all, the cases where we have sufficient (or even decisive) reasons for non-voluntary or voluntary euthanasia would be fewer than Finnis might fear. And if we

[85] For his views on this issue see Finnis (2011f, pp. 287-288, 311-312; 2011g, pp. 304-306).

[86] Finnis (2011d, pp. 222, 241).

descended from ideal theory to the domain of everyday practical decision-making, where we must act in ignorance and under uncertainty, such cases would probably be fewer still. For example, on most occasions during our everyday practical reasoning we might have very strong reasons to defer to the prediction each person does about the value of their future life, assuming that their judgment is more likely than not to be true than ours[87]. Of course, this is not tantamount to a rebuttal of Finnis objection, for surely he finds this implication only slightly less unsavoury. But then again, it is an objection that lacks bite against those who do not find implausible that some cases of non-voluntary or involuntary euthanasia are justified. I, for one, am one of them.

Thus, Finnis's version of personal Anthropocentrism is grounded on the false premise that genetic constitution has normative significance and, as is usual with this family of views, has implications I find very hard to accept. We can do better, as I shall try to show, in developing a view that is both contrary to foeticide and that claims that what really matters is the personal value of our lives. But we will just have to do without the claim that the life, in a biological sense, of a human animal is always personally valuable.

§5. *On agent-relative Anthropocentrism. Some final remarks on this family of views as a whole.*

So as to finish my review of the various anthropocentric positions on the wrongness of killing beings like us, allow me to consider a different approach to the issue. Thus far I have assumed that, be it personally or impersonally, human animals were *agent-neutrally* valuable. All rational agents are assigned the common aim of trying and preventing the death of human animals. It could

[87] Singer (2011, pp. 167-178) discusses in greater detail the morality of euthanasia and the problems posed by its non-voluntary and involuntary sorts.

be asserted, however, that being one of our kind has some kind of *agent-relative* value. It could be claimed

(5) It is always disvaluable that there is some being, *p*, and some state of affairs, *s*, such that *s* entails *p*'s killing a member of her own species.

and

It is good in that it gives *p* a reason to try and prevent *s* from obtaining.

Since we are human animals, this implies that our main reasons, usually decisive, not to kill other people are given by our sharing this attribute of species membership[88]. I will attack this position by way of a common variation of the imaginary scenario in which an extraterrestrial being comes to pay us a visit. So here she comes again, though now it is she who wonders, as a rational agent, what her reasons are, say, against killing normal adult human animals. If this agent-relative version of Anthropocentrism were true, then she would have to conclude that the strongest reasons she can have not to do it are absent. How should she treat normal human adults? One could say that she has as few reasons not to kill them as many people believe we have not to kill non-human animals. It would be justified for her to kill normal human adults for food or for clothing. If that sounds implausible and we want to claim that her reasons not to kill them would be very strong, then her reasons must be provided by some other facts. Though that would undermine claim (5) as a sound account of our reasons against killing, (5) might still be true. It might be that comembership in a species is somewhat agent-relatively valuable, so that our imagined extraterrestrial would have an extra reason to prefer, caeteris paribus, the survival of a member of her race to the survival of a human animal. Nevertheless, I cannot see how comembership in a species could amount to little more than that.

[88] This version of Anthropocentrism is discussed in Tooley (1983, p. 85) and Boonin (2003, pp. 26-27), but in special detail in McMahan (2002, pp. 217-228).

All in all, the main problem of the different versions of Anthropocentrism is their characteristic defence of species membership as a tremendously important attribute. As I said, its historical success is quite understandable. It gives us a clear-cut distinction between those conscious beings that matter most (human animals) and those others that matter much, much less (the rest). Also, since most humans are rational and self-conscious creatures, it is a proxy property which does a good job of tracking other properties whose normative significance is more plausible. As a non-ideal practical principle, it might have been justified to follow some version of Anthropocentrism when the requirement of acting on more sophisticated patterns of moral deliberation would have been simply too demanding.

That, however, is no longer so. I think I have shown how there is no way to account for the immense importance species membership has been given. Also, I have explained how adopting Anthropocentrism leads to counterintuitive conclusions regarding our reasons to commit suicide or to assist others in doing so, as well as regarding the killing of anencephalic infants or people in a PVS. I have also addressed the worry that abandoning Anthropocentrism may have terrible repercussions. Take severely cognitively impaired human animals, whose psychological capacities are similar to those of beings we treat in ways we would consider abhorrent to inflict on our species co-members. Does it mean, as our extraterrestrial friend reflected, that we may hunt those humans, or process them for food or clothing? If by rejecting Anthropocentrism we were bound to conclude such things, then we would have serious reasons not to do so. As I said, fortunately, that is not the case. We may recognise that our treatment of non-human animals has been very deficient and that the kind of life an animal with relatively low psychological capacities may live gives us reasons strong enough to make it usually decisively wrong to kill it, or produce clothes out of it or to experiment with it. If that is so, abandoning the relevance of species membership does not necessarily have unacceptable implications.

§6. *On the personal value of the contents of a life*.

In this section I will try to explain what I believe to be the most defensible view contrary to foeticide. Later on I shall try to persuade you that this is indeed the case, but let me provide you now with a hint of the reasons why. I am talking about Don Marquis's account[89]. Marquis's simple assertion could be put thusly:

> We always have strong reasons, usually decisive, not to kill an individual with a future like ours.

This position, unlike all the impersonal views I have sketched, manages to accommodate our intuition that whenever we have strong reasons against killing, it is usually because death is *bad for* the one who dies. And unlike Anthropocentrism it manages to avoid the mistake of attributing normative significance to species membership. It rests on an account of the wrongness of killing beings like us which, if not completely incontestable, is quite plausible once one accepts certain ontological and normative assumptions, namely, that we are essentially living organisms[90] and that *diachronic identity* matters[91].

Let me tell you something more about identity. Imagine a living organism, say, a tiger. It is hiding behind some bushes so that the middle section of its body is out of your sight –you can only perceive its head and its tail. How can you tell that head and tail are part of the same animal organism? Why not suppose that they are proper parts of two different organisms? Or parts of no organism? In the next section I shall say something more about what being an organism consists in, but for now we can rely on our everyday criteria for individuating

[89] The original formulation of his argument can be found in Marquis (1989). Over the years he has further developed and defended it in Marquis (1998; 2001; 2002; 2003; 2004; 2005; 2006; 2007a). This account seems to be also endorsed sometimes by Lee (2010, pp. 22, 64).

[90] Marquis (2001 abstract; 2003 p.439; 2007a:p.399; 2007b: p.195).

[91] This assumption underlies the original (1989) and subsequent reformulations of his argument (see references above), but is made explicit in Marquis (2003: p.439-440).

animals. You have seen tigers in pictures, in film, perhaps even in a zoo or safari. You know how they look like. You know how the sort of thing we call a tiger behaves. More generally, you know how multicellular animal organisms behave. Because of that you can tell there is one tiger behind the bush and that the head and tail you see are *spatial parts* of that organism.

That is what we call a judgement of *synchronic identity* –the claim that one or more chunks of matter occupying different spatial locations at a certain time are actually proper parts of a single individual entity. Suppose that, the tiger still behind the bushes, we set its tail on fire. That would certainly be bad for the tiger, since it would make it suffer. But imagine that, surprisingly, after setting the tail on fire we discover that there was a second tiger behind the bushes, to which the tail actually belonged. The suffering brought about by the burning of the tail could not now count as bad for our first tiger (the tiger to which the head belongs) but to the one whose existence we just learned about. Because of how tigers are, in order to determine for which one it is personally bad that the tail is burnt we need first to know to which one it belongs.

I guess what I just explained is pretty commonsensical. At least it keeps track of our practices of identification of the proper parts of large organisms, such as tigers or human animals. Now, some other identity judgements we make in our everyday lives are *diachronic* –instead of assigning different spatial parts to the same individual, they have to do with tracking an individual across time. That is what our father, Peter, has to do every school day. He drops Clara at school in the morning and picks her up in the afternoon. In the meantime, Peter has been hard at work in his office. For several hours he has lost track of her daughter, not being able to follow her movements through space and time. And yet, when the moment comes of carrying her back home, he seems to have no trouble in identifying her among the crowd of raucous children.

How can that be? Peter would most probably be baffled if we asked him how he knows Clara has persisted over time and how he knows that the girl he is

picking up now is the same as the one he parted with in the morning. Perhaps he would appeal to his memory of what her daughter looked like in the morning or to a recent photograph of her. Most of the time such criterion may get the answer right, but we know it cannot be the truth about what is required for Clara to persist over time. After all, the passage of time can make a person unrecognisable. Suppose that many years afterwards, during a scientific expedition in which she is participating, Clara's boat sinks, and she is marooned in a deserted island for several decades. After that time, she manages to return to civilisation. Surely Peter would not want to claim that the woman who claims to be her daughter must be lying because she looks nothing like Clara did when she was lost.

Some philosophers argue that we are essentially living organisms and that an organism persists through time while its component parts are trapped in the same metabolism. As I mentioned, I shall tell you more about that later[92]. Now, what it means is that an organism at t_1 and an organism at t_2 are one and the same just in case we can trace the history of a single metabolism at every temporal point between t_1 and t_2 in which all the parts that compose the organism that exists at that point are integrated. For example, assume that, unbeknown to Peter, Clara had split into two different organisms during school time, just like we know our cells can do. This is not something that happens to animals like us, but is common among other simpler living things, like amoebas. According to the conditions of temporal persistence of organisms I just sketched, an organism ceases to exist when there no longer is a single metabolism coordinating its parts. That can happen because its metabolic processes stop, causing the organism to die. But it can also happen after a process in which an organism divides, so that it is replaced by two others, each with its own metabolic processes. Thus, in this scenario, Clara would have ceased to exist, though by dividing instead of by dying, as it is usual. No matter which of Clara's descendants Peter would have picked up, or how much it resembled her, it is not her daughter whom he would have brought back home.

[92] We shall explore this view in greater detail in CH. III, **§2**.

According to Marquis, this account of what we are essentially and how we persist through time tells us something about how to determine whether something that will be good or bad for someone in the future will be good or bad for a specific person that now exists. Peter wonders whether Clara will be successful in her career. He just has to check whether any of those who will be in the future successful in their careers is the same organism as Clara, in case she lives. If Clara were to suffer from an incurable illness that would bring about her untimely death at age ten then, according to Marquis, we would be justified in claiming that she has been deprived of the good of being a successful professional. Death would have been bad for her at least in this respect. But suppose that we know that, had she lived, Clara would also have had a lovable partner and friends, would have develop quite a taste in architecture and photography and would have otherwise enjoyed very pleasant experiences. Of course, we know there would have been some bad things in Clara's life, but after these are aggregated to the good ones it turns out her life would have been very worth living. Hence, death has deprived her of all that. Assuming that the bad of death is somehow commensurate with the net sum of what she has lost, then dying has been bad for her indeed.

Succinctly put, Marquis's account would be[93]

Future-of-Value Account:

(a) The fact that dying would be bad for the one who dies gives us reasons not to kill her.

(b) The badness of death for the one who dies at a certain time is determined by the net value her life would have were she not to die at that time.

[93] Marquis (1989: pp.189-190; 2001 abstract; 2005 abstract; 2006 abstract; 2007a: p.399; 2007b:p.196).

(c) One can *only* be deprived, in the normatively significant sense, of a future that is *her own*.

How is this view contrary to foeticide? Let us assume that foetuses are living organisms of the same sort we are. That means that the future of the foetus consists of everything that will be good or bad for the child and adult that will develop out of it. If its life will be worth living, enjoying those goods we recognise as making the life of human animals more valuable than that of other beings, then we will have very strong reasons against killing it, usually decisive. Of course, some foetuses cannot be expected to have such bright future. It will even be true of some that their lives are worth ending. But then, that is also true of adults –some people have lives more worth living than other people have, and for some the net value of their life is negative, so that death would not be bad for them. Sometimes, then, there are just weak reasons against foeticide and, some other times, even reasons *for* it.

This account of the personal badness of death and the wrongness of killing seems to be safe from objections I made against other views contrary to foeticide. It appeals to facts which can hardly be denied to be reason-giving, such as how good or bad a life is for the one who lives it[94]. Anthropocentrism, instead, insists on implausible significance of species membership. In addition, unlike the authors who favour a personal version of Anthropocentrism, Marquis need not suppose that the mere fact of a human organism's being alive has any significance in determining what one of us loses in dying. Instead, our main

[94] Sometimes I'll speak of 'the personal value a life has for the one whose life it is' (or 'personal value', for short), but other times of 'the prudential value of a life'. I mean the same by both phrases.

reasons not to kill someone are given by the degree its future life, the life it will live in case it does not die, will have value[95].

Now, this account applies to all conscious beings, all those beings for whom there is anything that can be good or bad. Not all such beings have futures of the same value, so the strength of our reasons not to kill them will vary accordingly. Picture both the most fortunate of sardines and the most fortunate of dogs. Provided that what we know about the psychological capacities of sardines and of dogs is true, because of the worthwhile experiences this dog has in store our reasons against killing it given by the personal value of its life are stronger than those given by how good the sardine's life will be for it. Additionally, reasons of this sort that we may have against killing a dog with a pitiful, though still worthwhile, life will be weaker than those against killing the fortunate dog.

Regarding human animals, that is also true. Here we need an understanding of how much worthwhile a life must be in order for us to have reasons against ending it strong enough to account for all the cases about which we believe such reasons would be decisive. One worry we may have has to do with human animals whose psychological capacities are very limited. They will never be able to experience some of the goods that make most of our lives distinctively valuable. Does this mean that it is permissible to treat them in ways most of us would consider abhorrent? For example, it may turn out that we have sufficient reasons to kill them for food or clothing. As I mentioned earlier in this chapter,

[95] The identification of the value of a future is problematic in three ways. First, it can be doubted that death can be bad for us, since by dying we cease to exist. Second, it is not clear which of the possible futures an individual may have we ought to choose and it is unclear how the goods and bads in a possible future are to be aggregated. Second, even if we knew how to do these things we would still have to derive principles for assessing the value of possible futures in our non-ideal conditions of practical deliberation. Unfortunately, I will not be able to explore these issues. However, since (as shall be seen in CH. III, **§4** and **§5**) most views favourable to foeticide face at least some of these problems as well it will not be a factor affecting the relative plausibility of Marquis's position.

this worry presupposes that these are justified ways of treating non-human animals with similarly limited psychological capacities. I do not think that is true. We can, thus, avoid this unacceptable implication by admitting that our practical attitudes towards the death of many non-human animals should be similar to those we now have regarding cognitively impaired humans.

How does this account fare regarding those instances of lives not worth living? One unproblematic case, I believe, is that of organisms of our kind which (presumably) lack the structures necessary for consciousness (like anencephalic infants). Since they are *unapt* for attribution of personal value to their life, this account correctly implies that we have no personal reasons of this sort against killing them. It is also unproblematic in that it avoids the implication that we have decisive reasons never to commit suicide or to assist someone to do so even when her life would not be worth living. Yet its implication that the fact that someone's life is *worth ending* gives us reasons to kill her might be more troublesome. Does this mean that miserable people can be killed against their will? As I suggested when commenting on this worry as expressed by Finnis, one could try to appease those fears by recognising the value of autonomy or by suggesting that this is compatible with the policy of deferring in our everyday practical reasoning to the forecast each person does about the value of their future life. As to the value of autonomy, let me add here a distinction which can greatly benefit someone who endorses a theory such as Marquis's. We could distinguish between lives which are *irremediably* worth ending and those which are not. Regarding the former, it is to be noted that Marquis's account does not imply that our *only* reasons against killing are given by the personal value of someone's life. Thus, it is compatible with a view that attaches value (personal or impersonal) to autonomy so that it turns out that we have very strong reasons not to kill someone against her will. Moreover, as to those beings whose future life is *remediably* not worth living, it does not necessarily imply that we have strong reasons to kill them. Because this account attaches great importance to the personal value of lives, it is

compatible with the claim that we would have strong reasons to *try and improve* the situation of those beings so that their future becomes expectably valuable.

Let me end this section by commenting on an objection sometimes pressed against Marquis's account. Some[96] have argued that it suffers from the same problem I attributed to the impersonal version of Anthropocentrism, namely, that it implies that our reasons to produce a being like us are, at any time, as strong as our reasons, at that time, not to kill a being like us. I hope this presentation of the argument, as based on the personal value of lives, shows how they might have been wrong. For one, Marquis's argument need not assume that it is good *simpliciter* that there exist beings with lives worth living. Also, it must be remarked that at the time one is considering whether to create someone or not, no one exists of whom it is true that being created will be good for her. This is, then, one final respect in which Marquis's view fares better than the other alternatives which I have considered.

§7. *Concluding remarks*.

It might be objected that I have not yet considered a question which figures prominently in the discussion about foeticide. This is the debate about at what point in foetal development a human animal comes into existence. Anthropocentrism stresses the importance of being a human animal. Marquis's account assumes that we are essentially living organisms. This belief is also shared by many authors of an anthropocentric persuasion. Thus, on these views, we need to determine when human animals begin to exist in order to know when our strongest reasons against killing obtain. So as not to clot my main argument excessively, I tackle this question in APPENDIX A.

[96] Such as, for example, Norcross (1990), Reiman (1996), Savulescu (2002), Korcz (2002, pp. 582-585; 2004).

Now in this chapter I have reviewed what I take to be the most important views contrary to foeticide. My aim has been to find the strongest such position, which in this context means the one with fewer problems than the alternatives. In this my motivation is, if you will, wholly self-regarding. For my argument that we have decisive reasons to believe some view favourable to foeticide will work only insofar it is indeed true that our reasons to adopt a view contrary to it are not stronger. And, of course, the way to prove this is by comparing such favourable views with the best argument against killing foetuses.

I have argued that Marquis's is the best such view in the literature. First, by being based on the net value of the contents of a being's existence, it gives normative significance to what actually has it and avoids many implausible implications. Second, it can be complemented to overcome specific objections of its own. This is a view one can have relatively strong reasons to accept, provided that one buys its deeper assumptions. As I remarked, one of them is the normative significance of personal identity, which justifies the claim that someone loses some future good by dying just in case that good would have otherwise been *hers*. The other is the ontological claim that we are essentially living organisms. As I will try to show in the next chapter, however, the *Future-of-Value Account* is prone to many powerful objections.

CHAPTER III
VIEWS FAVOURABLE TO FOETICIDE

§1. *On the aim of this chapter.*

In this chapter I will elaborate on four powerful objections to Marquis's argument against foeticide. These objections have the peculiarity that, if true, they change the *Future-of-Value Account* into an argument for a view favourable to foeticide, at least for the greater part of a pregnancy. When I first introduced it I defined the *Future-of-Value Account* as comprising the following three claims:

(a) The fact that dying would be bad for the one who dies gives us our main reasons not to kill her;

(b) The badness of death for the one who dies at a certain time is determined by the net value her life would have were she not to die at that time;

(c) One can *only* be deprived by death, in the normatively significant sense, of a future that is *her own*.

As I remarked, in order to adopt a view contrary to foeticide Marquis needs to add the following premise to his argument:

(d) We are essentially human animals.

The first of the objections I will present merely targets Marquis's ontological assumption (d). It does so by denying that we are essentially human animals

and claiming, instead, that we are some kind of psychological entity, such as a mind or a person. As I argue in APPENDIX B, we ought to assume that the foetus acquires a mind around the eighteenth week after conception. Prior to that time it would be impossible that the foetus is numerically identical with the future person it will become. As a consequence, it would be false that before that time we have reasons against killing it given by the personal value of the life of the person it would eventually become.

The other three objections aim at disproving key elements of Marquis's account of the wrongness of killing. Consider premise (c). It assumes that diachronic identity is the *only* normatively relevant relation that can hold between an individual and the stuff that constitutes her future. From now on I shall express this contention as claim

(e) Identity is what matters in survival.

I believe that we ought to deny (e) and propose, instead, that what matters is the obtainment of some psychological relation. Because the foetus cannot have psychological properties before its eighteenth week, it cannot be, prior to that time, related to its future in the way that matters. Moreover, this psychological relation need not be constrained by the formal properties of numerical identity. In particular, it is ostensibly *gradual* in nature. This allows for the possibility that the personal value of a future varies with the intensity with which we are psychologically related to it. Since even a foetus older than eighteen weeks is so weakly psychologically related to its future, it can be argued that our most important reasons against killing it are correspondingly weak.

Thirdly, I will discuss an attack against premise (b). It consists in adding a further requirement for something to be valuable for someone by proposing that it cannot be in an individual's interest that something be the case unless that thing's being the case figures somehow as a content of a desire of hers. I shall express this contention as claim

(e) Prudential value is desire-based.

Since desires are mental states, a pre-eighteenth week foetus cannot entertain them. Hence, it cannot be against its interests to cease to exist. Different theories disagree about whether after that time it is correct to claim that, in a significant sense, a foetus desires to live. They also disagree about the strength of the reasons against killing it with which such desires provide us.

Finally, I will consider an objection to premise (a). This premise makes the wrongness of killing depend on the how bad dying was for the victim. This dependence can be challenged if it is argued, instead, that killing a person is very seriously wrong insofar it constitutes an act of disrespect towards an autonomous choice to continue to live. Assuming that autonomy is a psychological capacity, these grounds against killing cannot give us reasons against foeticide before the eighteenth week after conception.

It must be noted, however, that some of these objections may also entail a view favourable to infanticide. By this I mean that they may ground the claim that our reasons against infanticide are, caeteris paribus, *much weaker* than the reasons we have against killing individuals like you or me and that, as a consequence, the requirements for the justification of killing an infant are *much less* demanding. Many people find this implication deeply unsettling. Consequently, a complete defence of these positions will require an argument for the revision of our attitudes toward infanticide. My final task in this chapter will therefore be to analyse some strategies advanced to the end that the required revision is acceptable.

§2. *On denying that we are essentially human animals*.

Let us start by examining the ontological assumption that we are essentially human animals, as well as its denial. Allow me, first, to elaborate briefly on what precisely is being here discussed. Individual objects can be ascribed at least two different kinds of predicates[97]. On the one hand they can satisfy predicates expressing an *adjectival* or *characterising* concept. Thus, for instance, we may say that Clara is fair-haired and obedient, whereas Pounce is black-haired and obstinate. The concepts these predicate words stand for merely involve the specification of the property or properties an object must bear for the concept to be applicable to it, say, possessing hair of a certain colour or certain dispositions of character. On the other hand, individuals can satisfy predicates which stand for *sortal* concepts. Thus, for instance, we may say that Clara is a girl, whereas Pounce is a cat. These sortal concepts also involve a specification of the properties an object must bear for the concept to be applicable to it, say, being a woman of such and such age or being relevantly similar to certain exemplars. Yet, in addition, they provide an answer to the question about *what* something is[98].

---

[97] For this distinction among kinds of predicates I follow Strawson (1959, pp. 167-173), Lowe (2003, pp. 89-80) and, above all, Wiggins (2001), from whom I borrow the details of the explanation and some examples. A discussion of the substance sortal/phase sortal distinction can be found in pp.29-30; a summary glimpse of Wiggins's own thesis on the sortal dependency of identity and individuation (which he calls **D**) can be obtained from the core formulation of **D** (p.56 supplemented by p.60, first paragraph, *in fine*) and its ten glosses **D**(i) to (x), occurring at pp.64, 70, 72, 74 and 96.

[98] That is not to deny that adjectival predicate words may be part of a construction expressing a sortal concept. Thus, in 'Clara is a fair-haired girl' or 'Pounce is an obstinate cat'. But notice that here *being fair-haired* operates as a restriction to the sortal *girl*, just as *obstinate* restricts the extension of *cat*. And it is only because the sortal concepts furnish us with adequate principles of individuation and continuity that we can identify the objects for which we are to check whether the concepts of being fair-haired or obstinate apply. In general, the ability to apply an adjectival concept to an object presupposes the ability of singling out such object as falling under some sortal concept and ascertaining which changes in properties it can and cannot endure (see

The question about the sortal concept under which something falls is closely connected with the questions about the thing's identity and its continuity. For whenever we make a claim like 'Clara is a girl' we are implying also that there are girls and that Clara is identical to some girl. Moreover, if *girl* is a good covering concept for the thing that is Clara and if our grasp of such concept is adequate, it must furnish us both with an adequate principle of individuation and an adequate principle of continuity for Clara. That is, it must allow us both to single out Clara as a girl, distinguishing her from other girls and from other objects which are not girls, and to trace her movement through space and time, allowing us to determine at any given time, whether there exists a girl identical with Clara or whether, on the contrary, she has ceased to exist *as a girl*.

Because of the concept the sortal predicate 'girl' stands for, we know it is possible for an object to cease to be a girl (for instance, because she has grown older) and yet to continue to exist. That is because the concept of girl is of the kind sometimes called *phase* concepts. These are sortal concepts that need not apply to an object, in the present tense, at every moment throughout its existence. Other examples of phase concepts would be *tadpole* or *professor*. For we know it is possible for one and the same thing to stop being a tadpole and continue to exist as an adult frog, just as it was possible for Clara to leave girlhood behind and enter maturity. For any object, however, there is some sortal concept such that it applies to the object, in the present tense, at every time the object exists. These *substance* concepts, like all sortals, provide us with principles of individuation and continuity for the objects that belong to their extension. But, unlike phase sortals, they allow us to utter an unrestricted judgement of non-existence in the face of failure in the application of the relevant criterion of individuation or of continuity. For suppose that the substance sortal of beings like us is indeed *human animal*. If, as it has been

---

Strawson 1959, p. 168). Certainly, though, in many contexts of communication, the relevant sortal predicate can be left unuttered without loss in sense, since it is presupposed, and talk may deal with adjectival predication only (e.g., 'This is boring', during the screening of a film).

noted, there was at some time no girl identical with Clara, it would only follow that by that time Clara had ceased to exist as a girl. If, however, there was at some time no human animal identical with Clara, it would necessarily follow that Clara had ceased to exist *simpliciter*.

Hence, when it is claimed –as in (d)– that we are essentially human animals, what is asserted is that *human animal* is our substance concept and that it is to the principles furnished by this concept that we need to turn in order to keep track of beings like us and to distinguish those changes in properties that we can survive from those that result in our destruction. Now, none of the alternative positions that shall be discussed hereafter need to deny that we are somehow related to a human animal. They just need to deny that this is the kind of relation fixed by the fact that the concept of human animal is our *substance* sortal. Consequently, where those who claim we are essentially human animals argue that, of necessity, we share the fate of the human animal we are so closely associated with, its detractors are wont to show that it would be possible for us to part ways with such animal and that, therefore, it cannot be what we essentially are.

Let us follow the literature in calling the view that we are essentially human animals *Animalism*. What I shall do here now is to present the imagined situation (the *Brain Transplant* scenario) which prompted many philosophers –which I shall refer to in the customarily way as *Lockeans*– to claim that our persistence conditions are psychological in nature and incompatible with our being essentially human animals[99]. Then the general question will be raised

[99] Prominent Animalists include W. R. Carter (1982), David DeGrazia (2005, pp. 11-76), Eric T. Olson (1997a; 1997b), Paul Snowdon (1990) and Peter van Inwagen (1990). This position is different from the claim that we are bodies (and that, therefore, we continue to exist after death *as corpses*) pressed on, for instance, by Judith Jarvis Thomson (1997). Prominent Lockeans include H.P. Grice (1941) –arguably the one who initiated the contemporary debate on personal identity–, David Lewis (1983), Jeff McMahan (2002, pp. 66-94), Mark Johnston (1987), Derek Parfit (1971; 1984; 2008; 2012), John Perry (2008 [1975]), Sydney Shoemaker (1963; 1970; 1984; 1997) and Peter Unger (1990; 2000). For their part, Robert Nozick (1981) and David

whether there is any plausible alternative account of what we are compatible with such persistence conditions. I will start by developing the three main problems highlighted by Animalists which suggest a negative answer to that question –the *Too Many Thinkers* and *Too Many Persons Problems* and the *Epistemic Problem*. Finally, I will review those psychological approaches to what we are that seem able to surmount said objections.

The *Brain Transplant* situation is Sydney Shoemaker's influential variation of Locke's case of the prince and the cobbler exchanging souls[100]. Shoemaker invites us to envisage a future in which surgeons are able to successfully remove brain tumours by extracting people's brains, operating on them and then putting them back in their heads. He then describes the following event:

> "...a surgeon discovers that an assistant has made a horrible mistake. Two men, a Mr. Brown and a Mr. Robinson, had been operated on for brain tumors, and brain extractions had been performed on both of them. At the end of the operations, however, the assistant inadvertently put Brown's brain in Robinson's head, and Robinson's brain in Brown's head. One of these men immediately dies, but the other, the one with Robinson's body and Brown's brain, eventually regains consciousness. Let us call the latter "Brownson". Upon regaining consciousness Brownson exhibits great shock and surprise at the appearance of his body. Then, upon seeing Brown's body, he exclaims incredulously "That's me lying there!" Pointing to himself he says "This isn't my body; the one over there is!" When asked

---

Wiggins (2001, pp. 193-244) are friends of neither view; and though I hesitate to name him an Animalist, Bernard Williams (1973a; 1973b) was certainly contrary to Lockeanism. Eric T. Olson has excellent reviews of the competing answers to both this ontological question (2007) and the related question about our persistence (2010).

[100] For the prince and the cobbler imaginary case see Locke's *Essay concerning Human Understanding*, II, Ch. 27, Sec. 15: "[...] for should the soul of a prince, carrying with it the consciousness of the prince's past life, enter and inform the body of a cobbler, as soon as deserted by his own soul, very one sees he would be the same person with the prince, accountable only for the princes actions [...]"; I quote from Locke (2008, p. 44).

> his name he automatically replies “Brown”. He recognizes Brown’s wife and family (whom Robinson had never met), and is able to describe in detail events in Brown’s life, always describing them as events in his own life. Of Robinson’s past life he evidences no knowledge at all. Over a period of time he is observed to display all of the personality traits, mannerisms, interests, likes and dislikes, and so on that had previously characterized Brown, and to act and talk in ways alien to the old Robinson.”[101]

Now, Shoemaker feels “rather strongly inclined” to claim that Brownson is actually Brown, in spite of having Robinson’s body[102]. Many people share this intuition, though not all of them are happy with it[103]. Some of those who are, and have devoted their efforts to this matter, have attempted to flesh out, in different ways, the psychological criterion of personal identity over time that this intuition suggests. They have done so by drawing inspiration from John Locke, who distinguishes human animals (“men”, as he calls them) from the persons with whom they are related and famously defines the concept of a person as

> “a thinking intelligent being, that has reason and reflection, and can consider itself as itself, the same thinking thing, in different times and places”[104]

And shortly thereafter he claims that personal identity consists just in

> “the sameness of a rational being; and as far as this consciousness can be extended backwards to any past action or thought, so far reaches the identity of that person”[105]

---

[101] Shoemaker (1963, pp. 23-24).

[102] Shoemaker *ibid.*, p.24.

[103] This is because some of those who believe that Brownson is Brown at the same time adhere to Animalism, and must defend an account of the persistence conditions of animals that accommodates that belief.

[104] I quote from Locke (2008, p. 39).

Now, there are two distinct families of views, *narrow* and *wide*, in the contemporary Lockean tradition[106]. They disagree regarding the necessity of physical persistence for our continued existence over time. Thus, on the

> *Narrow View* the persistence over time of beings like us consists in non-branching psychological persistence caused by the physical persistence of the material stuff where our mental states or capacities are realised[107].

Allow me a note on the 'non-branching' (or *uniqueness*) requirement. This is necessary in order to maintain the transitivity of numerical identity, which any successful analysis of it needs preserve, lest it runs afoul of Leibniz's Law regarding the indiscernibility of identicals. This states that for any particulars, *x* and *y*, if *x* is numerically identical with *y* then, of necessity, any attribute possessed by *x* is also possessed by *y*, and vice versa. But this follows from the fact that identity is reflexive (can only hold between an object and itself), and the absurdity of the notion that an object's history be different from that self-same object's history. But suppose that we dropped the non-branching requirement in order to see what ensues. Assume we analyse *x*'s persistence as P-continuity.

---

[105] Locke, *ibid*. Here "consciousness" ought to be construed as *self*-consciousness, I believe, for Locke explains it as "perceiving that [one] does perceive" and knowing that "we see, hear, smell, taste, feel, meditate, or will anything", whenever we do so.

[106] I borrow the terms *narrow* and *wide* from Parfit (1984, p. 207). The reader will notice my *Wide* has been defined so as to include Parfit's *Wide* (continuity by a reliable cause) and *Widest* (continuity by any cause), since such a distinction is unnecessary for my argument. I do think my *Narrow* and *Wide Views* refer to the same positions, respectively, as Olson's *conservative* (2007, p. 17) and *liberal* views (pp. 19-20).

[107] In ordinary circumstances, of course, that material stuff is that which composes certain regions of a brain. What counts, then, as the persistence of a brain? Some authors allow that our brains can survive a process of sufficiently gradual replacement of its matter –Unger (1990, pp. 123-125), McMahan (2002, pp. 68, 70-71). Also, Unger imagines cases in which a sentient being comes to gradually exchange its brain for a new one (1990, pp. 156-159) or for some inorganic bionic unit in such a way that sameness of the material realiser of the being's psychology is nevertheless preserved (1990, p. 150; 2000, pp. 342-343).

Then a scenario is conceivable in which some future and simultaneously existing *y* and *z* are both P-continuous with *x*, and to the same degree (if P-continuity admitted of degrees). Thus, both *y* and *z* would have all it takes to count as *x*. Moreover, since identity is transitive, that would imply that *y* is the same particular as *z*. But that cannot be, for if we have correctly individuated the simultaneously existing entities *y* and *z*, it is possible that their histories diverge, and they are not numerically identical. Hence, P-continuity cannot establish identity when it branches. Or, alternatively, perhaps we were wrong in individuating *y* as distinct from *z*. But then only one future object is P-continuous with *x*, so that P-continuity does not branch. Anyhow, for an analysis of transtemporal identity to ensure that the formal properties of identity carry over to the analysans, something like the non-branching clause must figure in it[108].

Being that clarified, let us return our attention to the *Narrow View*. On this view, our intuition that Brown persists as Brownson can be quite straightforwardly explained. Brown's psychological features ostensibly continue in Brownson and, since the brain is preserved intact and merely suffers change in location, the kind of physical continuity required on the *Narrow View* obtains. There even is more physical continuity than it is strictly needed, for Brownson would still be Brown, according to this view, even if only Brown's *cerebrum* had been transplanted into Robinson's head. That is because the cerebrum –the newest, largest, walnut-like portion of our brain– seems to be that part on which our most distinctive mental states and capacities depend, whereas the posterior part of our brains, or *brainstem*, seems to be related to the vegetative functions of our organism.

On different versions of this view, however, different construals of the condition of psychological persistence are offered. Hence, for some the basis of

[108] I do not mean by this that the non-branching clause ought to figure explicitly in the description of some persistence conditions. Perhaps it suffices that the specific persistence conditions chosen exclude branching out of natural necessity.

psychological persistence is continuity of *mental content*[109]. This is typically specified in terms of psychological continuity and connectedness, which serve to account for the belief that persistence of our mental life requires that its changes be gradual and that present mental content –such as memories, beliefs, desires or character traits– be explainable by reference to past mental content and in turn serve to explain future ones[110]. One begins by elucidating the notion of a *direct psychological connection* using some examples: a memory is connected to an earlier experience; an intention, to the later act that carries it out; particular beliefs or desires entertained at some time, to, e.g., those self-same beliefs or desires present at a later time[111]. Since from time to time there

[109] Thus, Parfit (2008, p. 177; 2012, pp. 6-7), when he claims: (a) that only when non-branching psychological continuity caused by possession of enough of the same brain obtains between me and some future person that person is me; (b) that if neither kind of persistence obtains between me and some future person, that person is not me; and (c) that if one of the two kinds of persistence obtains between me and some future person but the other does not, the question whether that future person is me has no answer, even though there would be nothing that we did not know. For his part, Shoemaker (1997) defends the view that persons are substances whose persistence requires psychological continuity. But he also adopts the view that *particular* mental states are identical with *particular* physical states of the brain or the nervous system, and concludes that psychological continuity necessarily involves physical continuity (pp. 293, 297-298). Johnston also endorses the *Narrow View* (1987, p. 64), for he denies that the particular embodiment of our mind is contingent and requires persistence of the same brain alongside psychological continuity.

[110] Compare with Nozick's elaboration of the notion of continuity (1981, p. 35): "[t]o say that something is a continuer of x is not merely to say its properties are qualitatively the same as x's, or resemble them. Rather it is to say they grow out of x's properties, are causally produced by them, are to be explained by x's having had its properties, and so forth. [...] The later temporal stages [of something] also must be causally dependent (in an appropriate way) to the earlier ones". Unfortunately I cannot afford to discuss Nozick's account of our identity through time (the closer continuer schema). In a nutshell it holds that "y at t_2 is the same person as x at t_1 only if, first, y's properties at t_2 stem from, grow out of, are causally dependent on x's properties at t_1 and, second, there is no other z at t_2 that stands in a closer (or as close) relationship to x at t_1 than y at t_2 does" (1981, p. 37). Both physical and psychological properties (including bodily continuity) must be aggregated in order to judge who someone's closest continuer is. Hence, Nozick is neither a Lockean nor an Animalist.

[111] I borrow these examples from Parfit (1984, p. 205).

can be more or less direct psychological connections, psychological connectedness is a *gradual* relation. It must, however, be noticed that since psychological connectedness is *not* a transitive relation, whereas numerical identity is, the latter cannot be analysed in terms of the former. For suppose a series of individuals *a* to *z* such that *a* exists at some time and each member of the series exists at a later time than its preceding one. Suppose further that between any two contiguous members of the series, *a*-*b*, *b*-*c*, *c*-*d*... *y*-*z*, there is a great amount of direct psychological connections. This is compatible with there being no psychological connections between *a* and *z*. Thus, it is possible that we ought to judge (1) that *a* is the same as *b*, which is the same as *c*, which is same as *d*... and *y* the same as *z*, but (2) that *a* is numerically distinct from *z*. Such, possibility, however, is excluded by the transitivity of identity. Thus, identity cannot be analysed solely in terms of psychological connectedness.

We can, fortunately, construct a transitive relation out of psychological connectedness –the one called *psychological continuity*. Suppose that *a* is directly psychologically connected both with *b* and with *c*; *b* with *c* and *d*, and so on. Then, we can say that, even though *z* is not directly psychologically connected with *a*, it is linked to it by an *overlapping chain* of psychological connections. Whenever this happens we say that *a* and *z* are psychologically continuous, so that we preserve the sense that, even though a person's mental life at some later moment in her life may bear little resemblance with her mental life as it was at some earlier moment, it may nevertheless be the case that her later mental states have grown out of her earlier ones[112]. On this version of the

[112] Here I have offered the broadest account of the relations of psychological connectedness and psychological continuity. They can be specified in several different ways. For example, Parfit defines psychological continuity as the holding of overlapping chains of strong connectedness, and then goes on to claim that there is strong connectedness when "the number of direct connections, over any day, is at least half the number that hold, over every day, in the lives of nearly every actual person" (1984, p. 206). See especially the discussion of Parfit's view in McMahan (2002, pp. 39-66), which, though centred on these psychological relations as grounds for prudential concern rather than identity, explores some of these possibilities.

Narrow View, then, we have reason to claim that Brownson is Brown because the former is psychologically continuous with the latter, where this continuity is caused by their sharing the same brain.

On a second version of the *Narrow View* continuity of mental content is unimportant for psychological persistence, the preferred construal being sameness of the relevant psychological *capacities*, such as the capacity for consciousness. Thus, the functional continuity of the brain regions responsible for those capacities is enough[113]. In this way, Brownson would still be Brown even if he was unable to recall any event in Brown's life from a first-person perspective, could offer no answer to the question about his name, recognised none of his relatives and acquaintances and otherwise lacked all of the psychological states, such as beliefs and desires, which were distinctive of Brown –simply because he would have inherited Brown's functioning brain.

So much for this family of Lockean views. Now on the

> *Wide View* the persistence over time of beings like us consists in non-branching psychological persistence, even if it is not caused by the physical persistence of the material stuff where our mental states or capacities are realised.

On this view then, we might survive the destruction of our entire body, including our brain, provided that there is one future person with whom we are

[113] Thus, Unger, who believes that X at t_1 and Y at t_2 are the same person just in case "sufficiently continuous physical realization" of X's basic psychological capacities obtains (in a non-branching way) between the physical realiser of X's psychology and the physical realiser of Y's, allowing that a certain amount of loss of such basic capacities can be compensated by the preservation of enough of the person's other central non-basic psychological capacities and central mental content (1990, pp. 140-143). McMahan endorses a similarly flavoured view, on which the criterion of personal identity is "the continued existence and functioning, in nonbranching form, of enough of the same brain to be capable of generating consciousness or mental activity" (2002, p. 68).

psychologically continuous[114]. This is the philosophical view that makes sense of the notion that teletransportation is just the fastest way of travelling, and not some way to die[115]. In these cases, as usually described, our body is scanned and then disintegrated in a booth here on Earth and subsequently replicated atom by atom in, say, a booth on Mars. The resulting person is an exact physical duplicate of the one who entered the booth on Earth. Since the brain's structure has also been replicated, the resulting person is also a perfect psychological duplicate. Thus, the person who exited the booth on Mars and the

[114] Thus, Grice (1941, pp. 340-344), who claims that persons or selves are logical constructions analysable in terms of a series of successive "total temporal states" (sets of simultaneous experiences occurring at a given time) linked by chains of event memory, and takes this to be "mainly a modification of Locke's theory" (p.340). Shoemaker (1984, pp. 108-111), for his part, suggests a case of survival without physical continuity involving a "brain-state transfer" procedure. This would consist in recording a person's total brain-state, destroying her body and then restructuring the brain of a duplicate body in order to make its state qualitatively undistinguishable from the recorded brain-state. Psychological continuity is then, for Shoemaker, sufficient for our persistence over time (1970, p. 278 n.18). As to Derek Parfit, he was also of the opinion at one point (1984, p. 208) that sameness of brain was neither necessary nor sufficient for our persistence, but he came later to see this as a mistake, which he corrected in subsequent reprints of his *Reasons and Persons* (see my footnote 109 *supra*). Regarding John Perry, he admits that actual cases of memory involve some sort of physical continuity, but he seems to concede that that need not be so in some imaginary cases (2008 [1975], p. 148). Since he believes that we should incorporate into our account of personal identity mental content other than memory (2008, p. 154), perhaps he would be ready to generalise this concession. Finally, David Lewis (1983, pp. 55-56) requires that psychological continuity and connectedness be brought about in such a way that the "character" of "each succeeding mental state" is causally dependent on the states "immediately before it". Lewis takes this requirement to exclude those cases of apparent regular succession of mental states actually caused by accident or by an agent that creates "a succession of mental states patterned to counterfeit our ordinary mental life". It seems to me, then, that, on his own description, the sort of causal continuity he requires is compatible with physical discontinuity of whatever stuff our mental life is realised in. All we would need is some reliable causal mechanism to fill in the discontinuity, like Shoemaker's brain-state transfer device. Finally, Noonan declares that "what is crucial for personal identity is neither identity of body nor of brain, but psychological continuity [...]" (2003, p. 214).

[115] The teletransportation thought-experiment was first introduced by Parfit (1984, Part Three).

one who entered the booth on Earth are as psychologically continuous with one another as Brownson was with Brown.

All of these Lockean views about our identity over time present a challenge to a view contrary to foeticide such as Marquis's. That is because on any of them we could not have begun to exist before the foetus acquires a mind, and it is very unlikely that this happens before the eighteenth week after fertilisation. My reasons for this claim can be found in APPENDIX B. In this way, on every one of these views, killing an early foetus does not qualify as destroying one of us, but merely as preventing one of us from existing. From a moral point of view early foeticide is much more akin to contraception than to homicide. Regarding the late foetus, different construals of psychological persistence yield different results. If this condition of persistence is understood in terms of sameness of the capacity for consciousness, or if psychological continuity is taken to require fairly weak psychological connectedness over time, then it can be defended that the late foetus is the same being as the future person it will become. A revised version of Marquis's position could be true, holding only for foeticides committed late in pregnancy. If, however, a more stringent conception of psychological continuity is demanded, the foetus can only count, no matter its stage of development, as our predecessor, but not as numerically identical with us. Killing a foetus would be at any time in pregnancy much like contraception. Marquis's account of the wrongness of killing could not support, then, a view contrary to foeticide.

As I advanced, though, this Lockean approach to our persistence has been vehemently contested. It has been argued that we are essentially an animal and that the persistence conditions of animals are incompatible with those proposed by Lockeans[116]. The first step in the formulation of the persistence conditions of animals, conceived of as living organisms, is to understand what a *life* is. Peter

[116] This is to be distinguished from the claim, advanced by some Lockeans, that the persistence conditions of minded non-human animals are also psychologically based (Unger, 2000, pp. 337-343).

van Inwagen imagines how some disembodied intellect would describe its first observed living organism:

> "What I am observing is an unimaginably complex self-maintaining storm of atoms. This storm moves across the surface of the world, drawing swirls and clots of atoms into it and expelling others, always maintaining its overall structure. One might call it a homeodynamic event."[117] [i.e., a self-moving, self-maintaining event]

A life would be, then, an event constituted by the sum of activities of some particles of matter[118]. Of course not any activities will do[119]. As the disembodied intellect noticed, it must be the case that these particles succeed in maintaining their structure in spite of being constantly exchanging matter and energy with their surroundings. It must also be the case that these metabolic activities (as they are called) are self-directed, rather than being regulated by the environment. Moreover, these self-directedness must respond to some internal plan (in the case of the organisms with which we are acquainted, this plan is chemically based). Whenever a collection of particles of matter is engaged in this sort of activities, they compose an organism[120]. This suggests an analysis of the persistence of organisms in terms of persistence of a life. Thus, we can specify for animals, including human animals, a

---

[117] Van Inwagen (1990, p. 87). Here I will be following Peter van Inwagen's account of the ontology of living organisms (1990, pp. 81-97), subscribed also by Olson (1997b, pp. 126-142; 2007, pp. 27-29), and very similar to DeGrazia's (2005, p. 245). In fact, as van Inwagen (1990, p. 143) and Olson (1997b, p. 137; 2007, p. 28) explicitly acknowledge, their answer to the problem of animal identity was anticipated by Locke: "an animal is a living organized body: and consequently the same animal, as we have observed, is the same continued life communicated to different particles of matter, as they happen successively to be united to that organized living body" (Locke, 2008, p. 37). Locke added that "in an animal the fitness of the organization, and the motion wherein life consists, begin together, the motion coming from within" (p. 36).

[118] Van Inwagen (1990, p. 89).

[119] Olson (1997b, pp. 126-131).

[120] Van Inwagen (1990, p. 90).

> *Biological Criterion*, i.e., for any *x* existing at t_1 and any *y* existing at a later time t_2, where *x* and *y* are animals, *x* and *y* are the same animal just in case the activities of the particles that constitute *y*'s life at t_2 are causally continuous in the appropriate way with the activities of the particles that constitute *x*'s life at t_1.[121]

If we are identical with the human animal with which we are so closely associated, then our persistence conditions are those spelled out in the *Biological Criterion*. We began to exist when the group of cells we refer to as 'zygote' became sufficiently integrated to constitute a life[122]. We shall cease to exist when that integration becomes impossible to maintain and death ensues. This also entails that we continue to exist as long as our metabolic activities carry on in a self-regulated way, even if we happen to lose our psychological capacities –that is, being a person is just a possible phase in the life of a human animal, consisting in its being conscious of itself as a temporally extended entity.

A few pages ago I pointed out that some Animalists do believe that Brownson is Brown. One may wonder how they accommodate that intuition with their further belief that the persistence of a human animal requires appropriate causal continuity of its metabolism and other vital functions. Suppose you have your finger cut. It is no longer part of your organism, since your organism is composed of the material particles caught up in the activities that constitute your life and the particles of matter arranged as what was your finger are no longer caught up in such activities. Moreover, having your finger cut in no way troubles the normal functioning of your organism. Your brainstem, which is the region where the structures relevant for the regulation of your vital activities

[121] See van Inwagen (1990, pp. 145, 148-149) and Olson (1997b, pp. 135, 138). Of course the 'in the appropriate way' clause is obscenely vague. I shall specify it for some cases (those needed in the argument) but not for any conceivable problematic situation. Such cases as, e.g., those of frozen organisms, fission of an organism, embryo development and metamorphosis are discussed by van Inwagen (1990, pp. 145-158).

[122] As discussed in APPENDIX A.

reside, has been in no way affected by the loss of your finger. Neither have the various systems (cardiovascular, respiratory, digestive, etc.) responsible for those activities. A similar story could be told in case you had your whole hand cut. The severed hand is no longer a part of your organism, and in no way troubles its continued existence. That would be true also if you lost an arm, or both, or all of your extremities.

Could the mutilation be more drastic? Can a human animal survive as a severed head? The contention is that if the brainstem continues to be attached to the severed head there can be appropriate causal continuity of the organism's vital functions, since the structures in which they are realised persist. I say 'can be' because, since the organism has lost the organs that were responsible for the performance of those functions, it will be in need of artificial aid. If this aid is provided, at no time the material particles that compose the head will stop being caught up in self-regulating, internally directed, metabolic processes[123]. But then there is no difficulty to admitting that a headless brain is a living organism. If the necessary artificial aid is provided, then the brainstem will continue regulating the processes that keep the cells that compose the organism working together. Admittedly, it has seen the matter it imposed such processes on diminished. Just as it was diminished, but in a lesser degree, when a finger was cut off. Just as it would be augmented if the brain received a new body. The belief that Brownson is Brown can, thus, be accounted for in the following way. First, the organism named Brown is mutilated until it is reduced to a brain. Then, someone moves it to another spatial location and appropriately attaches it to a brainless body. Finally, the

[123] See van Inwagen (1990, pp. 169-181), Olson (1997b, pp. 44-46). But what is the fate of the cells that once composed the organism minus the head? According to van Inwagen, they compose no living organism, not even if it were provided with an artificial brainstem so that its organs continued functioning and decay was avoided, since the information that would direct its activities would not "be produced by the action of parts of an object that [the cells] are also parts of, but by a machine, that is, by simples which are caught up in the life of no organism" (p. 179).

organism sees the number of material particles over which it exerts its control augmented, until they coincide with the body it was attached to[124].

You may remember that I claimed that on any Lockean view Brownson would still be Brown even if only the cerebrum had been transplanted, provided psychological persistence also obtained. For obvious reasons, Animalists will deny this[125]. Suppose that it was only Robinson's cerebrum that was irreparably damaged, so that its brainstem is not removed. That means that the metabolic activities which constituted Robinson's life carry on unimpeded and regulated from the brainstem, assuming proper nutritional support is provided. In short, what we have on the operating table is Robinson himself, very much alive, but lacking part of his brain. A similar tale can be told about Brown when his cerebrum is removed and readied for the transplant. What remains lying on that operating table is not Brown's former body, but Brown himself, deprived of his capacity for thought. When what was Brown's cerebrum is transplanted into Robinson's body it ends up trapped in the metabolic activities regulated from the brainstem to which it is now attached. Brownson is just the animal we called Robinson with a new seat of consciousness.

---

[124] As I pointed out at the beginning of the previous paragraph not all Animalists will concede that Brownson is Brown. One who doesn't is DeGrazia (2005, pp. 142-149) on account of endorsing a less stringent conception than van Inwagen's about what it takes for an event to qualify as a life. The evidence seems to suggest that maintenance of the circulatory-respiratory function is possible without a brainstem, and DeGrazia believes this is integration enough. But this is another aspect of the debate in APPENDIX A, and I don't think this disagreement augments the vulnerability of the Animalist position. Though not an Animalist (or a Lockean either), it is worth noting that Wiggins (2001, pp. 232-236) doesn't share the intuition that Brownson is Brown. According to his Locke-inspired *forensic conception* of personhood, there must be room also for non-psychological predicates in our account of what a person is – how she "stands or walks or frowns or smiles or laughs or sulks or earnestly entreats, or how he fries an egg" (p. 234) and her "physiognomy" and "physical presence" (p. 135). All these modes of activity which we could truly attribute to Brown are not truly attributable to Brownson. This would justify the identity judgment that Brownson is not Brown.

[125] Van Inwagen (1990, p. 179 n.63), Olson (1997b, pp. 111-119; 2007, pp. 41-42).

This may sound counterintuitive in that it seems to fit ill with the practical attitudes we would expect to have as Brown, Brownson and their relatives and acquaintances. When we picture ourselves as undergoing the same ordeal as Brown, we might find that whether our whole brain or just our cerebrum is transplanted into Robinson's body makes little difference regarding what we feel about the situation. Both from a first-person and third-person perspective the experience will be similar: we would be able to recall Brown's life, as if it were, from within, offer the right answer to the question about Brown's name, recognise Brown's relatives and acquaintances and otherwise possess all of the psychological states, such as beliefs and desires, which were distinctive of Brown.

The claim that the Animalist story of what happens in this cerebrum transplant fits ill with our practical attitudes, however, presupposes that different metaphysical tales force us to different practical attitudes towards this case. But the Animalist can deny this by refusing to accept the claim that identity has the significance in practical reason that this objection assumes. But this entails denying what I have called claim (e) –i.e., that identity is what matters in survival. Discussion of this point belongs in the next section, but let us here simply say that the Animalist and the Lockean, in spite of their ontological disagreement, *can* agree that the relations underpinning psychological persistence are the ones that actually have normative significance, whereas numerical identity (conceived as it may) strictly speaking has none or very little. An appeal to our practical attitudes in cases of cerebrum transplants such as this is, then, no good objection.

Anyhow, the Lockean will claim that Brownson is Brown, and not Robinson with Brown's cerebrum, and that the other cerebrum-less body in this story is not Brown, though it was Brown's body once. Yet the Lockean can hardly deny that what was Brown's body is a living organism, and an animal, and human, for the same reason it would be strange for her to claim that an anencephalic infant, besides not being a person, is not a human animal either. Where did that animal

come from? There are two possible answers. It either began to exist when Brown's cerebrum was removed, or it existed before that and continued to exist after the cerebrum's removal. But the first alternative is implausible for it implies three implausible contentions. First, that the removal of cerebra is a procedure for introducing new animals into the world. Second, that while the person exists there is no animal associated with it. Third, that the foetus that became me ceased to exist when its cerebrum became active and, consequently, that the activation of cerebra which are part of animals is a successful procedure for removing animals from the world. Surely, the second alternative constitutes a more parsimonious answer. The animal on that operating table is the one who had always accompanied Brown. It had preceded Brown as an unconscious foetus, Brown made use of it for a while and now it has parted ways with him.

But if that is true of Brown, then it is true of all persons we are acquainted with. Thus, wherever we stand there are a person and an animal, and both Animalists and Lockeans agree that they are very closely associated. They share many of their attributes. For instance, if it is true that Clara is one and a half meters tall, it must be because that is an attribute of her human animal. If she is in China, then her human animal is too, and so on.

This is the root of the first ontological problem presented to the Lockean views – the *Too Many Thinkers Problem*[126]. It would be the following: this sharing of properties seems to lead to baffling conclusions regarding the population of thinkers in the world. Suppose Clara is in pain. Now we can point to the physical structures in her brain and nervous system where pain is realised. But those structures belong to the human animal too, so it must also be in pain –in terms of its physical states it is indistinguishable from you, so it must also be psychologically indistinguishable from you. Thus, there are two distinct entities in pain, Clara and her human animal, where common sense dictated that there was only one. Similar things can be said about other mental states, such as beliefs and desires. Whenever one of us entertains a thought, our animal is

[126] These problems for the Lockean views are pointed at by Olson (2007, pp. 29-37; 2010).

thinking it too (or, perhaps, thinking another, qualitatively indiscernible *token* though). In this way, the Lockean view multiplies the amount of thinkers that inhabit the world –twice the number of persons.

The trouble, though, runs deeper than that –a *Too Many Persons Problem* arises. If it is true that any psychological feature a person possesses is also possessed by the human animal associated with her, then the human animal must also qualify as a person in the Lockean definition of the term. For, if we are thinking intelligent beings able to perceive ourselves as temporally extended, then our animals are too. This means that wherever we are there are two persons, ourselves and our animal. And these two are qualitatively indistinguishable, except, strangely enough, in their persistence conditions. That is because one of the two (us) is *essentially* a person, not an animal, and therefore has psychological persistence conditions, while the other is essentially an animal, and only contingently a person, and therefore has biological persistence conditions. Accordingly, (1) the Lockean views double the number of persons in existence; (2) are wrong in claiming that all persons have the same, psychologically-based persistence conditions (non-animal persons do, but animal persons do not); (3) must explain why qualitatively indistinguishable entities have different persistence conditions (and, thus, belong to difference substance sortals).

Finally, it is claimed that if the Lockean views are right, then we face the *Epistemic Problem* of being unable to know whether we are the non-animal person or the animal one. For suppose that we think 'I am the non-animal person'. The truth-value of the proposition that is the meaning of that sentence depends on the reference of the personal pronoun 'I'. But this is an indexical term which always refers to the one who thinks it or utters it. So the proposition at issue will be true if thought by the non-animal person, but false if thought by the animal one. Whenever we think it, whose thought is it? Is it the non-animal's or the animal's? The Animalist contention is that there no way to for us to identify the correct answer. For all we know we might be wrong whenever we

think that we are the non-animal person and that we could survive the fantastic situations imagined by Lockean ontologists[127].

If the Lockean views are to be an acceptable alternative to Animalism, a relationship between ourselves and our human animal must be specified such that these problems are solved or shown to be no problems after all. Indeed, if our metaphysics is good, it ought to give us the means to count the objects that inhabit the world. Were it true that Lockean personal ontology fails in this regard, it would miss its theoretical aim. First, I will consider two ways of arguing for a distinction (that between derivative and non-derivative properties) which may help us solve the counting problems posed by the *Too Many Thinkers* and *Too Many Persons Problems*[128]. Secondly, I will consider how these strategies may overcome the *Epistemic Problem*.

Instead of claiming that we are *identical with* some human animal, Lockeans may claim that we are *constituted by* one[129]. For the sake of tradition, let us begin our cursory examination of constitution by contemplating the relation

[127] In all fairness, Olson (2007, pp. 219-222) concedes that Animalists must also find some way of specifying the relationship between a human animal and the *chunk of matter* it is made of that avoids the multiplication of thinkers, persons and the epistemic problem. Furthermore, Dean Zimmerman (2003) has a lengthy argument/review of the literature according to which these are problems every plausible materialist account of our ontology must deal with.

[128] Other strategies may consist in denying the real existence of one of the two things. Thus, we can deny that 'person' expresses a substance concept and retreat to the position that it is merely a phase sortal –once, Parfit (2008, pp. 203-205) suggested this move. Alternatively, we can deny that animals exist: the material particles which are vulgarly said to compose animals actually compose nothing, and thus the existence of animals is equated to the kind of virtual or nominal existence which some philosophers believe artefacts have. But the first is a fall-back position to be adopted in absence of any other viable alternatives. And my desire to keep the metaphysics of my argument as commonsensical as possible counsels me against denying the existence of animals if I can avoid it: "Organisms look like paradigm cases of material things. To deny their existence is to deny the reality of all ordinary objects" (Olson, 2007, p. 219).

[129] I shall make use of Lynne Rudder Baker's account of constitution as it appears in (1997; 1999a; 1999b; 2002a) since it is the one that has been discussed in the literature.

between a statue, say, Michelangelo's *David*, and the piece of marble that makes it up. Both entities, the statue and the piece of marble, are spatiotemporally coincident. Also, they share many of their other properties, such as height, weight, shape and so forth. Let us focus on their weight. Certainly, if *David* weighs n kilos, then the piece of marble weighs n kilos. But it would be a mistake to infer from this that it makes sense to add up their weights and conclude that the combined weight of David *and* the piece of marble is 2n kilos. When confronted with this problem of addition, we might feel inclined to follow the strategy favoured by Animalists in the discussion about personal ontology and simply say that *David* is numerically identical with the piece of marble that makes it up. Numerical identity helps us here because it is a *unity* relation. Whenever we say, for example, that *x* is numerically identical with *y*, where *x* and *y* are persons, we do not mean to say that there are two persons, but just one. This is because *x* and *y* are one and the same thing, so that there is only one thing exemplifying the attribute of being a person. Thus with the attribute of weighing n kilos. There are not two n-weighters in our story: *David* and the piece of marble are one and the same thing. Hence, there is just one exemplifier of the attribute of weighing n kilos.

Yet the claim that a statute and the piece of marble that makes it up are identical has a consequence that may be found undesirable, namely, that they have the same persistence conditions. For this claim entails, for example, that Michelangelo could not have chosen to sculpt *David* out of a numerically distinct (but perhaps indistinguishable in its intrinsic properties) piece of marble. Had Michelangelo so chosen, *David* would never have existed. Certainly there would be some statue to which everyone would refer as 'David'. This 'David' would be indistinguishable from our *David* in its intrinsic properties. It would be, however, another, distinct statue. Further oddities ensue, however. For it is not the case that, in this world where *David* is no statue, it does not exist. We assumed here that the statue *David* and its piece of marble were numerically identical. Since the piece of marble predated the statue, we can make sense of this by saying that 'being a statue' is a possible phase in the history of a piece of marble. We

cannot, then, refer to one and fail to refer to the other. The very same thing we refer to as *David* does exist in this world. It is the unsculptured piece of marble Michelangelo discarded.

Supposedly, the relation of constitution manages to avoid these unpalatable implications by enabling us to have it both ways[130]. It allows us to maintain at the same time that *David* and the piece of marble that makes it up (that 'constitutes' it) have different persistence conditions *and* that there are not two exemplifiers of being an n-weighter, but just one. Constitution occurs when some material thing belonging to a certain substance sortal, by gaining possession of certain properties in certain circumstances, comes to coincide with a new material thing belonging to a different substance sortal which possesses those properties essentially. Thus, for instance, when a statue emerges from a piece of marble if the latter is sculpted; or when a person emerges from a human animal if the latter acquires certain psychological capacities. So, certainly, constitution entails that there are two numerically distinct overlapping things. But consider again the weight of the statue. If *David* is n-weighty, it is so in virtue of being constituted by the particular piece of marble that makes it up. Had it been constituted by another piece of marble, its weight might have been different. Sculptures *borrow* their weight from whatever constitutes them. But the piece of marble would have weighted the same even if it had constituted no statue. In this respect, it is no borrower, but a genuine exemplifier. This suggests that the piece of marble is the primary or non-derivative bearer of the attribute of being an n-weighter, whereas the statue is its derivative bearer. Though the attribute of being an n-weighter is borne also by the statue, the piece of marble is its sole exemplifier. Again, if the piece of marble is so aesthetically pleasing, it is so in virtue of constituting a statue and not, say, a building block for an office skyscraper. The statue is aesthetically pleasing in a non-derivative way, whereas the piece of marble is so only derivatively.

[130] Here I draw heavily from Baker (1999a, pp. 156-160; 1999b, pp. 157-158; 2002a, pp. 373-375), also from Olson (2007, p. 65).

The relation of constitution is, then, one way to explain how an individual can have a certain property without exemplifying it itself, but merely by borrowing it from its actual exemplifier. If the Lockean view is true and if persons are constituted by human animals, then we can avoid overpopulating the world with thinkers and persons. We are thinkers and persons because we exemplify those attributes. The animal that constitutes us is a thinker and a person because it borrows this attributes. We would be thinkers and persons even if we came to be constituted by a numerically distinct animal. The animal, however, can only entertain thoughts, including the belief of being temporally extended, in virtue of constituting us. It is not a separate thinker, or a separate person. When we are careful to count persons and thinkers properly we see that we and our human animal are the *same* thinker and the *same* person, just like the statue is the *same* n-weighter as its piece of marble[131].

Some philosophers have their misgivings about constitution[132]. It is appropriate, then, to show how the distinction between derivative and non-derivative properties can be grounded in another relation, namely, the part-whole relation. According to the *Embodied Part View* we are a "the conscious and controlling

[131] Sometimes, it appears that Parfit believes that something like this is the right way of reducing the existence of persons to the existence of bodies and mental events: "Rather than claiming that there are no entities of some kind, Reductionists should distinguish kinds of entity, or ways of existing. When the existence of an X just consists in the existence of a Y, or Ys, though the X is distinct from the Y or Ys, it is not an independent or separately existing entity. Statues do not exist separately from the matter of which they are made. Nor do nations exist separately from their citizens and their territory" (Parfit, 1995, p. 18). He calls this view *Constitutive Reductionism*.

[132] Thus, Olson (2007, pp. 65-71) wonders whether constitution is yet another primitive relation and asks whether there is any principled way of distinguishing what he calls *constitution-inducing* properties (those properties which, when acquired by an object make it the case that it coincides which a new one) from those that are not constitution-inducing. For his part, McMahan (2002, p. 90) suggests that the relation of constitution is unintelligible, and that it collapses with numerical identity.

part" of an animal organism[133]. Human animals think, but do so only derivatively, that is, in virtue of having a part that does the thinking. When we are careful to count how many thinkers there are we do not take the animal and its thinking part to be separate thinkers

> "Animals digest their food by having a part, their stomach, that does the digesting. Animals sneeze by having a part, their nose, that does the sneezing. These facts do not create a Too Many Digesters or Too Many Sneezers Problem."[134]

Just as we are the primary subjects of mental acts, the animal we are a part of is the primary subject of many physical attributes we possess only derivatively. Thus, we refer to ourselves as having such attributes as height, weight, or having blond hair and hazel eyes. This linguistic phenomenon ought not to surprise us, supporters of this view claim. After all, we are a proper part of that animal and the function we perform for our organisms is such that they become our instruments of sustenance, perception and agency –it is only natural that we have come to view many attributes possessed by our organisms as *our own*.

This suggestion is related to a further one which may help us solve the *Epistemic Problem*[135]: it might be that our personal pronouns are ambiguous, so that we use them in different ways. Thus, suppose someone says

> 'I have been splashed'

But assume that actually only her trousers had been splashed. This fact need not make it false that she had been splashed. That is because 'I' is being used

[133] Different versions of this view are defended by McMahan (2002, pp. 92-94) and Parfit (2012, pp. 14-19), from whom I borrow the denomination.

[134] Parfit (2012, p. 15).

[135] Parfit (2012, pp. 20-24), from whom I borrow the splashing example which shortly follows.

widely in order to refer also to the clothes which that person was wearing. Similarly, the truth-value of

'I am the non-animal person'

Will depend on how its thinker was using the word 'I'. Following Parfit's proposal, let us use 'Inner-I' to refer to the part that does the thinking for the animal. And 'Outer-I' to refer to the human animal. Now we can disambiguate. If it was meant that Inner-I is the non-animal person, then the proposition is true. Moreover, we (the thinking part) know it to be true, for we understand the way we have chosen to use the pronoun. If, conversely, it was meant that Outer-I is the animal person, then the proposition is false. Moreover, the animal would know that it is false. For the animal only knows derivatively, and only what its thinking part (Inner-us) knows directly. Since its thinking part understands this disambiguation, the human animal understands it too[136].

It seems to me that, all things considered, the Lockean view is a tenable proposal regarding our persistence conditions. It is compatible with several approaches to the relation with the human animal with which we are associated which at least have a good chance of overcoming the problems highlighted by Animalist authors. All in all, I do not think we ought rationally to prefer Animalism. If that is true, then there is no need for us to accept the view contrary to foeticide that stems from Marquis's account of the wrongness of killing. This is significant because, as I have shown before, if we replace Marquis's ontological premise with a Lockean position, keeping the rest of his argument unchanged, a conclusion favourable to foeticide can be inferred from it.

[136] This proposed solution to the *Epistemic Problem* appears to be compatible as well with the view that we are constituted by a human animal. It might be more accurate in that case, though, to speak of 'Constituted-I' and 'Constituting-I' instead of 'Inner-I' and 'Outer-I'.

Someone who accepted the reasoning I have thus far followed could be, nevertheless, quite upset by the following. Suppose a Lockean view that analyses identity in terms of persistence of mental content. Provided that a sufficiently stringent conception of psychological continuity is demanded, not only shall foetuses count as numerically distinct from the persons they would become, but also newborn infants. Then, this modified *Future-of-Value Account* would yield a conclusion favourable to infanticide. I shall confront this problem in **§6** of this chapter.

§3. *On denying that identity is what matters*.

Now, there is a second, independent way of challenging Marquis's conclusion contrary to foeticide. As I said, one of the premises in his argument is

(c) One can *only* be deprived by death, in the normatively significant sense, of a future that is *her own*.

For this claim to be true it must be the case that identity is the *only* normatively relevant relation that can hold between an individual and the stuff that constitutes her future. In short, it must be true that

(e) Identity is what matters in survival.

But (e) can and has been denied, either by those who claim that identity has no normative significance, or those who claim that it has small normative significance. Whereas Animalists and Lockeans cannot agree on the *metaphysical* question about our survival, they can agree on the *practical* question about what matters in survival. That is because even if they disagree on what our identity over time consists in, they can agree that some psychological relation, and not personal identity, is the (main) basis or grounds of prudential concern. Thus, the relevant question to ask about some future

good or bad in order to determine what prudential reasons it gives us is not whether it will be *ours* but whether we shall be in the appropriate psychological relation with it[137]. It will help us understand how it can be that identity is not fundamentally significant if we consider some scenarios in succession[138].

Suppose that, when presented with the case where only Brown's cerebrum is transplanted into Brown's head, you were asked whether Brown has any grounds for being prudentially concerned about Brownson's future, that is, concerned about that future as if it were *his own*. As I mentioned, many philosophers agree that he does. This is a case in which, in all plausible specifications of the normatively significant psychological relations, these obtain between Brownson and Brown: the relevant brain regions persist and, with them, all of Brown's psychological capacities and content.

Let us change somewhat Brown's story. Suppose that several years before he needed any body-transplant Brown lost one of his cerebral hemispheres. Perhaps it died because of a stroke, or perhaps Brown had to have it surgically removed for medical reasons. Typically, our cerebral hemispheres are functionally asymmetrical, that is, each half of our cerebrum subserves different functions. It seems that usually the right half is specialised in language, whereas the left half processes spatial relations and controls emotions[139]. Thus, unless the other hemisphere adapts to the change, some of Brown's capacities

[137] Authors who believe that identity is not what matters include DeGrazia (2005, pp. 286-288; 2012, pp. 29-34), Holtug (2010, pp. 73-84), Kagan (2012, pp. 162-169), McMahan (2002, pp. 69-86), Parfit (1971; 1984, pp. 245-306; 1995) or Shoemaker (1970). Olson (1997b, pp. 52-72) believes at least that Animalism is compatible with the claim that numerical identity has no practical significance.

[138] I shall argue that we ought to believe that identity is not what matters. But that is just a way of claiming that we ought not to assume that the foetus has the same moral status as a person. Since that is Thomson's (1971) fundamental assumption, this explains why I shall neither be discussing her argumentative strategy, nor that of those who have sophisticated it over the years.

[139] Hugdahl (2005, p. 120).

will be diminished. If this happened, we would probably say that Brown, nevertheless, continues to exist. Let us, however, reinforce this claim by supposing that, unlike most people, Brown's cerebral hemispheres are functionally symmetrical. In this way, we can be sure that the loss of one half of his brain will not cause an impoverishment of Brown's capacities or mental life.

If we believe that what fundamentally matters in Brown's survival is preserved after having his cerebrum transplanted into Robinson's brain; if we also believe that it is preserved after losing a cerebral hemisphere, then surely we ought to believe that what fundamentally matters for Brown is preserved as well after an ordeal consisting in a combination of both situations. Consider the situation Jeff McMahan calls,

> *Loss Followed by Transplantation*. Brown suffers an accident that destroys one of his hemispheres and damages his brainstem and body. But, because Brown's hemispheres were largely symmetrically developed, virtually all of his cognitive capacities are preserved in the remaining hemisphere. Surgeons therefore extract Brown's undamaged hemisphere and transplant it into the cranium of his identical twin, Robinson, whose cerebrum has been destroyed but whose brainstem and body are intact[140].

Let us refer to the person with Robinson's body and Brown's sole hemisphere again as 'Brownson'. Lockeans will want to say that Brown has survived as Brownson. Animalists will insist that Brownson is numerically distinct from Brown. Yet anyhow Brown's first-person perspective, capacity for consciousness and mental life are all preserved in Brownson. It would be a mistake for Brown to feel sorry for his impending death but glad that a new life was in store for his twin. Quite the opposite is true. Before the transplant, Brown ought to be as prudentially concerned about his expected future (family and

[140] This particular description of the case (as well as that of *Division*, used below) appears in McMahan (2002, pp. 22-23), though I have reworded it so that it refers to Brown. Similar scenarios are considered in Parfit (1984, pp. 254-255).

friends, work and hobbies) as he was before. Also, he would have reasons to be prudentially concerned for experiences which would have otherwise been considered a part of Robinson's life. For suppose that, before the accident, Robinson had been suffering from an excruciating medical condition, such as kidney stones. It would be irrational for Brown, *before* the transplant, to be prudentially unconcerned for that condition since, *after* the transplant, it can be verily said, at least, that for all practical purposes it will be as if he were suffering from it.

In all the imaginary cases I have presented –the original Brown-Brownson scenario, its two variants and *Loss Followed by Transplantation*– only Animalists could claim both that Brownson was not Brown and that nevertheless the relation that matters the most for prudential concern obtained between Brown and Brownson. Lockeans had to admit that Brownson was Brown. Thus, it is not the case that from a Lockean perspective these cases challenge the pre-theoretical notion that identity is the grounds for prudential concern. In order to achieve that we need a further case

> *Division*. Brown is a member of a set of identical triplets, all of whom are involved in an accident. While Brown's brainstem and various vital organs are irreparably damaged, his cerebral hemispheres are unharmed. In the case of both other triplets, Robinson (1) and Robinson (2), however, their brainstems and bodies are undamaged but their cerebrums are destroyed. Surgeons are able to extract Brown's cerebrum intact but, instead of transplanting it whole, they divide it and transplant each hemisphere into the body of one of the two remaining triplets. Because Brown's hemispheres were symmetrically developed, the two people who are brought to consciousness after the operations –Brownson (1) and Brownson (2)– are both fully psychologically continuous with Brown as he was before the operation. Both believe themselves to be Brown and both have bodies almost indistinguishable from Brown's own.

As I explained in the previous section, any successful analysis of identity must ensure that the analysans preserves the formal properties of numerical identity. Since one of those is transitivity (i.e., that for any *x*, *y* and *z*, if *x*=*y* and *y*=*z*, then *x*=*z*), it must not be the case that our account of identity over time allows for two individuals to be the same and their history to possibly diverge. What must be the case is for such divergence to be necessarily excluded. *Division* illustrates why the 'non-branching' clause must be added in any psychological analysis of the persistence of persons. Suppose that a criterion of our identity over time merely specified some sort of psychological persistence. Given our answers to the imaginary cases thus far presented, we would be rationally compelled to admit that both Brownson (1) and Brownson (2) *are* Brown. This is because *each* of them has inherited Brown's psychological capacities and mental life by acquiring the regions of Brown's brain in which they were realised. It would be unjustified to say that *just one* of them is Brown while the other is not. Yet, since identity is transitive, we would be further compelled to admit that both Brownson (1) and Brownson (2) are the same person. But surely it is possible, e.g., that Brownson (1) is furiously working in his office while Brownson (2) is eating sushi during his vacation in Tokyo. Also, as time passes we can expect their beliefs and preferences to diverge to some extent, so that one may come to move in with Clara, considering her a wonderful choice for a life partner, while the other considers her a poor one. Is it possible that Brownson (1) and Brownson (2) are the same individual but that their histories (i.e., what they did last summer or whether they moved in with Clara) differ? We could, of course, consider the mereological sum of Brownson (1) and Brownson (2) as *one* person, yet spatially scattered and in possession of two separate fields of consciousness. This person could have these seemingly incompatible properties in virtue of having two different parts that possess them. Surely, though, given our concept of a person, it is highly implausible to consider such mereological sum as one[141]. If that is so, Brownson (1) and Brownson (2) are numerically distinct from one another and neither of them is Brown.

[141] The different ways in which we could describe Brown's survival as both Brownson (1) and (2) are considered and discarded as implausible in Parfit (1984, pp. 256-258).

Yet this is what is to be expected in those cases where the relations in which identity consists branch –none of the persons that result from the branching is identical with the original one. In the strictest metaphysical sense, by dividing Brown has ceased to exist. How should Brown feel when someone explains this to him right before his division? Well, in one way division is like death. There is no-one with whom the person is identical after *Division*, just as there is no-one one with whom she is identical after death. But in another way *Division* is like strict survival. When one survives, there is someone in the future who inherits one's psychological capacities and mental life by acquiring the regions of the brain in which they were realised. In *Division*, this is what happens also, only that there are two future individuals, instead of just one, who have all that would ordinarily take for that person to survive. Surely, if what it takes to survive obtains twice it would be irrational for Brown to consider his division as bad as death. In Parfit's words Brown would then be

> "... like someone who, when told of a drug that could double his years of life, regards the taking of this drug as death. The only difference in the case of division is that the extra years are to run concurrently. This is an interesting difference; but it cannot mean that there are *no* years to run. We might say: 'You will lose your identity. But there are different ways of doing this. Dying is one, dividing is another. To regard these as the same is to confuse two with zero. Double survival is not the same as ordinary survival. But this does not make it death. It is even less like death.'"[142]

We can further test our beliefs about whether *Division* would be like ordinary survival by supposing again that, before the accident, one of the Robinsons had been suffering from kidney stones. It would be irrational for Brown, *before* the transplant, to be prudentially unconcerned for that medical condition. No matter which of the Robinsons was afflicted by it, *after* the transplant it can be verily

[142] Parfit (1984, p. 262).

said that at least for all practical purposes it will be as if Brown were suffering from it[143].

The fact that division is most unlike death and the fact that we would be prudentially concerned for the fate of our future selves show that identity cannot be what solely matters in survival; also, if the intuitions stirred by this series of cases are to be relied upon, some psychological relation does matter[144]. Now, as I said, there are two ways of denying (e); in the

> *Weak way*, identity is not what fundamentally matters in survival, even though it has some *small* normative significance

Whereas in the

> *Strong way*, identity has *no* normative significance

Since Animalists analyse our persistence in terms of a biological relation, it would be implausible for one of them to adhere to the weak claim that identity has in itself some small normative significance. If one believes identity over time is analysable, one must accept that it borrows its importance from the relations

[143] It may be useful to note here how a fine metaphysician has expressed his perplexity towards the claim that there are some changes such that it is *as if* Brown can survive them. When asking himself whether "in some important sense" one may survive a situation like division, Peter van Inwagen (1990, p. 212) recognises that he doesn't understand this question if posed in strictly metaphysical terms. Although he immediately accepts that there is a non-metaphysical sense of survival (surviving "in [our] work", "in the memory of those who know [us]"), it is clear from the preceding discussion in his book (pp. 210-211) that he would regard his division as bad as death.

[144] David Lewis (1983) disagrees, and insists that there is another way to understand what happens in cases like *Division* that allows us to preserve the common-sense intuition that identity, analysed in terms of some psychological relation, is what matters. He manages this by embracing an ontology of temporal parts and by bestowing to the relations that prudentially matter the same formal properties as numerical identity –what a case of *Division* shows is that, previous to the fission, there had *always* been two persons sharing their temporal stages.

in which it consists, and I cannot see how the sole happening of facts about metabolic activities and the like can serve as grounds of prudential concern[145]. Animalists ought to claim, rather, that identity has *no* normative significance because the facts that account for it have none. In this sense, Animalism offers a strong basis on which to deny that identity matters[146]. Unlike Animalists, Lockeans can deny that identity is what fundamentally matters in either the strong or the weak way. On these views our identity over time is nothing more that some psychological relation or relations that hold in a non-branching way. Since they typically consider these relations to be normatively significant, their choice is whether and when the fact that they branch makes any difference. If it were conceded that the presence or absence of branching affects our prudential reasons, it would be admitted that whether numerical identity obtains or not has some importance[147].

There is *some* disagreement about which psychological relations ground prudential concern. Thus, some would say that only psychological continuity and connectedness (persistence of mental content) is important. Others would have that functional continuity (persistence of mental capacities) is also important, and even necessary. Moreover, there is additional disagreement about whether these relations only ground prudential concern when they are caused by physical continuity of those brain regions in which they are realised,

[145] Cf. Parfit (1995, pp. 28-33): this follows at least if we assume that relative to the lower-level facts on which Animalists ground their analysis, the higher level fact of identity is merely conceptual. Holtung (2010, p. 73) reaches the same conclusion. Marquis (2003, p. 439) suggests, on the contrary, that whilst we exist as beings without a very high degree of psychological unity ("five-year-olds", and, we can assume *a fortiori*, as non-minded beings) identity may link us to our futures in a more significant way that prudential concern. But how can the mere continuity of vital functions be normatively significant? He gives no argument in support of his claim.

[146] Olson (1997b, p. 71) is well aware of this, and Parfit (2012, p. 27) recognises as much.

[147] Parfit (1984, pp. 264-265) has a brief discussion about some cases in which branching may make a difference, either for the better or for the worse.

so that, for instance, teleportation would be about as bad as death[148]. I am not interested in discussing here the relevance of physical continuity, for there is no need. In the cases we encounter in our ordinary lives and, specifically, in the case of prenatal human psychological development, physical continuity causes, and thus correlates with, functional and psychological continuity.

Some the most influential theories in applied ethics resulting from these considerations about prudential concern are Jeff McMahan's time-relative accounts of both the personal value of death and some of our reasons against killing[149]. Let us grant for the sake of the argument that persistence of psychological capacities also grounds prudential concern –I shall assume that it matters that some future good (or bad) happens to someone who either shares my capacity for consciousness or whose capacity is descended from mine, e.g., by division. Whenever this relation of sameness or ancestry obtains, then I have prudential reasons to care for such future good. That is true even if I also know that there shall be little to none psychological continuity and connectedness between me now and the person who will enjoy it, perhaps because I know that by that time the person will be in the later stages of dementia. The question now is whether the presence of substantial psychological continuity and connectedness makes any difference to my prudential reasons. McMahan believes it does. Some fictional situations will help fix our intuitions on this matter. Consider first

[148] Famously, Derek Parfit believes that the relations that prudentially matter are psychological continuity and connectedness obtained by *any* cause (1984, p. 287). But, previously, Sydney Shoemaker had expressed similar thoughts (1970, pp. 284-285). David Lewis also finds that what he "mostly want[s] in wanting survival is that [his] mental life should flow on", even if he establishes some restrictions as to what counts as proper causes of continuity (1983, p. 55). Others such as McMahan (2002, p. 79) or Holtung (2010, pp. 85-128) believe that continuity of psychological capacities is also important and that the cause by which these psychological relations obtain must be compatible with the physical continuity of the material stuff in which our psychology is realised. Both DeGrazia (2005, pp. 286-288) and Steinbock (2011, pp. 76-79) explicitly adhere to McMahan's view about prudential concern.

[149] McMahan (2002, Ch. 2 and 3).

Valjean's life and Madeleine's life. There is a boy, Jean Valjean, who had a very hard life. He was born in a very poor family and soon became an orphan. Having to provide for her sister and her children, he dispensed with education and spent long hours every day working as a wood-cutter. Eventually, he was forced to resort to thievery and stole some food. He was imprisoned for twenty years, where he was frequently vexed and brutalised. During that time, he became self-centred and unsympathetic towards the welfare of other people.

There is a man, Monsieur Madeleine, whose mature years were full of accomplishments. He started a successful business in a small city which made him a very rich man, and yet he invested most of his benefits in several philanthropic initiatives. After some time, he became the city mayor, earning the respect of the citizens for his zeal and honesty. In his old age he took up a derelict woman's child as his stepdaughter, resigning both his private and public offices in order to care for her. He became a very loving father, being loved in return. One day, when he received news that the boy with whom his daughter was in love was in great peril, Madeleine set out and risked his life in order to save him. The three of them lived happily for a time, until finally Madeleine died.

Now,

Case One. After his death, some of Madeleine's documents make it apparent that he was, in fact, Jean Valjean. After being released from prison he was determined to become a better man and, being aware of the social prejudice against former convicts, he decided to change his name.

Case Two. Jean Valjean and Monsieur Madeleine were two different persons existing at the same time in different places who lived their lives unaware of each other. After being released from prison, Valjean's life was

dull and uneventful and he died at an old age. Monsieur Madeleine's early years had been unremarkable and unfulfilling.

Case Three[150]. After some time, it was discovered that Jean Valjean and Monsieur Madeleine were the same person, though the latter was unaware of this. Some scientist explained that after being released from prison Jean was severely traumatised by his past experience. So much so that he accepted the scientist's offer of a drug that would destroy all of his distinctive psychological features, including his beliefs and desires, as well as radically change his values and character traits. But it would also erase the memories of his time as a convict. After administering the drug, the scientist took care of Valjean for a while, encouraged him to choose a name, gave him some advice and sent him on his way.

How ought we to judge Valjean's life? Arguably, it depends on which version of the story is true. Suppose it is *Case One* that describes what actually happened. On the whole, I believe, it was a good life. In one respect, it would have been *better* if its first half had been absent, for it was full of misery. In another respect, Jean's life receives part of its overall value from the fact that it became so fulfilling *after* such initial hardships. Also, if it was to be represented in a graph accounting for the passage of time and the amount of good things present at each time, Jean's life as Madeleine would have resulted in an ascending line, culminating towards the end in an act of moral heroism. This pattern of constant improvement seems to account for part of the appeal of Valjean's life. Our judgment would have been very different, I believe, if events were to be rearranged so that he spent most of his later years in prison. And not just because such hardship would have been more difficult to endure, but also because it would have meant ending in a very poor note what had otherwise

[150] I borrow the fundamental features of this third scenario from McMahan's *The Cure* (2002, p. 77).

been a very rewarding existence. Thus, *what* happened in Jean Valjean's life matters for its overall value, but *when* it happened seems to matter as well[151].

Suppose now that *Case Two* was true. Valjean's life would merit a very different consideration. Depending on which is the right assessment of his latter dull, uneventful years, it ought to be deemed, I suggest, either barely worth living or not worth living at all. Yet we are well aware that at the same time Valjean experienced all that misery, elsewhere Madeleine was carrying on a life full of accomplishments. We might even believe that Madeleine's life in *Case Two* is, on the whole, more worth living than Jean Valjean's life was in *Case One*. Nevertheless, we know it would be wrong to claim that these facts about Madeleine's life somehow affect how personally valuable Jean Valjean's life is for Valjean himself in this case. We know why. Madeleine is not related to Valjean in any way that matters for prudential concern –for one, we can count two separate consciousnesses or minds in this situation; also, they are not psychologically connected. Thus, the goods and bads (and their relations) whose aggregation compose Madeleine's life provide Valjean with no prudential reasons.

But what about *Case Three*? One distinctive feature of this case, just like in *Case Two* and unlike *Case One*, is the radical psychological discontinuity between Jean Valjean and Monsieur Madeleine. McMahan's intuition when discussing cases like this one is that the degree in which a life is a single unit for prudential concern, as opposed to a mere aggregation of such units of varying lengths, depends on the degree of *psychological unity* within that life, which is a compound of psychological continuity and connectedness:

> "The degree of psychological unity within a life between times t_1 and t_2 is a function of the proportion of the mental life that is sustained over that

[151] Very interesting reflections about what affects the overall value of a life can be found in McMahan (2002, pp. 174-185) and Kagan (2012, pp. 241-281).

> period, the richness or density of that mental life, and the degree of internal reference among the various earlier and later mental stages"[152]

In *Case Three*, we can identify two clearly distinct units for prudential concern obtaining in succession in the same life, Valjean-Madeleine's. Hence, even if Valjean-Madeleine could somehow be aware of the extremely fulfilling life which awaited him after taking the drug, the rational attitude for him would be to regard those goods as if they were to be enjoyed by someone else. That is, the strength of the prudential reasons (and, thus, personal reasons) one has at a certain time to care about some event in one's future is a result of how good or bad such event is multiplied by the degree of psychological unity between oneself at that time and oneself at the time that good would be enjoyed. If the maximum degree of psychological unity obtains, then the multiplier is 1; as the degree of psychological unit decreases, so does the multiplier. As McMahan says, psychological unity functions as a "discount rate" for prudential concern[153].

Things would be very different if numerical identity was the relation that mattered. Since identity does not admit of degrees, then the strength of our prudential reasons to care for each of the goods or bads that constitute our life would not vary over time. In order to determine the personal value the obtainment of some future state of affairs has for some entity we would not need to relativise the problem for some time in the existence of said entity. That is because at all times in its existence the entity is related to the same degree with the future state of affairs at issue –we could, thus, treat prudential reasons as *time-neutral*. Identity, however, is not what matters for prudential concern, so the question about the prudential value which the obtainment of some future state of affairs has for an entity needs to be relativised to a particular time in its existence –thus the notion of *time-relative valuation* as a species of personal valuation, and of time-relative reasons as a species of personal reasons.

[152] McMahan (2002, p. 75).

[153] McMahan (2002, p. 80).

Since psychological unity functions as a discount rate, for those individuals who display a highly complex and integrated mental life the question about the time-relative value of their future life and the question about its time-neutral value shall have very similar answers. When that is not the case, such as in the life of our Valjean-Madeleine in *Case Three*, the answers can be very different. Madeleine's life has tremendous time-neutral value for Valjean right after he takes the drug; nevertheless, its time-relative value seems to be very little.

Suppose we accept, as the *Future-of-Value Account* assumes, that

(a) The badness of death for the one who dies at a certain time is determined by the net value her life would have were she not to die at that time;

Since this account further assumes

(c) One can *only* be deprived by death, in the normatively significant sense, of a future that is *her own*.

and accepts that identity is what fundamentally matters, 'the net value her life would have were she not to die at that time' clause has to be construed as stating that the personal value of death for some individual is a function of the difference between the *time-neutral value* her life would have if she were to die at some time and the time-neutral value her life would have if she were not to die at that time. Suppose we want to know how bad would be for Clara to die at t_1. Let us say that, before t_1, Clara's life had contained slightly more good than bad and that from t_1 to t_2 it will contain considerably more good than bad. It would be a considerable misfortune for Clara to die a t_1 rather than continuing to live until t_2. That is because if she dies, then her life would have been only slightly worth living, whereas if she continues to live until t_2 her life will be considerably worth living. But suppose now that we knew that a few months

before t_2 Clara would become afflicted with a medical condition that would inflict her terrible suffering, so much so that these few months would contain very little good. It would still be true that Clara's life until t_2 would contain a considerable surplus of good. Yet it would contain even more net positive value if she were to die right before she becomes afflicted with the disease. This gives us strong reasons to prefer that Clara dies at that time.

This seems to be in accord with the intuitions of most people. However, this is no reason to prefer a time-neutral account of the net value of our lives, for in those cases (such as Clara's) in which the deceased's life enjoyed a high degree of psychological unity it much resembles their time-relative value. We need to see what happens in those cases where each account provides us with distinct results. One of such cases is *Division*. Since this account assumes that identity matters, it wrongly tells us that what happens to our descendants cannot add to or detract from the prudential value of our lives. Another such case is Valjean-Madeleine's life. It wrongly tells us that what happens to Madeleine adds to the prudential value of Valjean's life in exactly the same way it adds to the prudential value of Madeleine's life. But we have seen how what fundamentally matters is not identity. We may, then, in a modified *Future-of-Value-Account* replace (c) with

(c') In the sense that matters, one is fundamentally deprived by death of those future events with which one stands in the appropriate psychological relations.

We know that these relations may vary in strength over time and that, consequently, the prudential value or disvalue death has for some individual at the time of dying is *time-relative:* we need to apply a discount rate for the psychological disunity existent between the person at the time of death and each of the goods or bads that constitute her future life (or that of her

descendants) at the time when they would accrue[154]. This not only gives us the right answer in such cases as *Division* and Valjean-Madeleine's life. It also accommodates some further widespread beliefs. For example, it tells us that, from a prudential perspective, caeteris paribus, the death of a person whose future life would contain a certain amount of net good is less preferable than the death of a dog whose future life would contain exactly the same amount of good. Since the relations that matter hold to a lesser degree between the dog and its future life, he loses in dying less time-relative value than the person does. It is worth noticing though, that from a time-neutral perspective the amount of good lost is the same[155].

[154] This is McMahan's own *Time-Relative Interest Account of the Badness of Death*, on which he constructs his *Time-Relative Interest Account of the Wrongness of Killing* conscious non-autonomous beings. He calls its time-neutral counterpart the *Life Comparative Account* (2002, pp. 105-106). McMahan recognises that this is the view of the badness of death that underlies Don Marquis's argument, and criticises it accordingly (2002, pp. 270-272).

[155] It must be stressed, though, that even if McMahan believes that in our moral deliberation the present interest in continuing to live of a minded being (i.e., its interest in continuing to live at the time of action) is to be considered time-relative, that is not true of all interests. One ought to distinguish here between merely *possible interests*, which are those 'the individual can be expected to acquire if one acts one way but not another' (2006, pp. 629-630), and *future interests*, which are those 'it is reasonable to expect the individual to have independently of one's present action' (2006, p. 629). In this respect, on the one hand, he admits that prudence may allow us to treat our *own* future interests as time-relative. But, on the other hand, he believes that morality requires 'that we not discount the future interests of others' for the weakness of their psychological relation with their future selves, but rather that we treat those interests as time-neutral (2006, pp. 629, 632 n.4). Thus, an individual's interest in not receiving an injury is to be treated as time-neutral, since it exists independently from the agent's causing or failing to cause the injury. But since an interest in a future state of affairs shall only exist at the time such state of affairs obtains provided that the individual is not killed, when considering whether to kill an individual such interests must be classified as merely possible, not as future ones. In this sense, McMahan claims that 'to evaluate [an individual's] loss by reference to how strong her interest would have been' at the time of the obtainment of the relevant state of affairs 'would be to assign the interest a weight it will never have'. Thus, our value judgment must be time relative, i.e., made by reference to how strong her interest was at the time the person is killed.

As to the wrongness of killing, we are assuming for the sake of our argument that our reasons against killing someone are given by the fact that death is bad for her. That being so, our reasons against killing foetuses (and infants) cannot be as strong as those we have against killing a person. Suppose again that we can determine that, if Clara were to die at a certain time, she would be deprived of many significant goods, which would greatly outweigh what bads were in store for her. Since a person displays a high degree of psychological unity, then we can say that the effect of the discount rate in determining how time-relatively valuable they are for her is non-negligible but weak, so that her future goods and bads contribute to the value of her life for most of what they are worth. Death would be very personally bad for Clara. If so, then it can be justifiably said, first, that she has very strong prudential reasons to desire not to die and to try and prevent her death from happening. Second, it can also be justifiably said that we would have correspondingly strong reasons to try and prevent her death from happening and, consequently, not to kill her.

On the other hand, suppose that a foetus (or newborn infant) were to die at a certain time. Suppose, further, that what the future held in store for it was that mixture of goods and bads which, were they to be events in our own life, we would consider it to be worth living –perhaps even more so than most people's. What reasons would we have to desire, *for the foetus's sake*, that its death does not ensue? If what mattered was the time-neutral value of the foetus's life, our reasons would be weighty indeed. That is Marquis's intuition, and the conclusion his argument leads us to. What matters, however, is the time-relative value of the foetus's life, which allows us to reach a very different conclusion. First, consider the early foetus, that is, the foetus before it has developed any psychological capacities[156]. On some views about what we are, it is numerically identical with the future person it will become; on some other views, it isn't. Thus, for some, early foeticide is strictly speaking killing one of us, whereas on others it is merely preventing one of us from existing. Yet from the point of view of practical reason, all may agree that it counts as preventing the existence of

[156] McMahan (2002, pp. 267-269).

one of us, as if one were employing a contraceptive method. When pondering whether to kill an early foetus, one is *only* choosing whether the world will be such that someone will own that good, but *not* choosing whether to deprive the foetus of it. Consider the future good that shall not exist in the world in case the early foetus dies. It is the same good which would be absent if that foetus had never been conceived. That future good, however, is related to no-one in the way that matters for prudential concern either when the foetus does not exist or whilst it exists prior to its obtaining any psychological capacities. For an early foetus, death is as bad as never to have existed –that is, not bad at all. If we have reasons to desire that an early foetus continues to live and reasons not to bring about its death, these cannot stem from concern about what is valuable for it[157].

[157] It is worth contrasting this view about the early foetus with the one defended by Elizabeth Harman (1999). She appears to acknowledge that whereas the death of a 'person' is bad because it is 'bad for the subject who died' (p.316) that cannot be true of the *early foetus*, i.e., the foetus before it has become a 'subject of experience' (p.310). Thus, she claims that a person has moral status whereas an early foetus doesn't. This can be explained because of her adherence to the position (which she explicitly attributes to Derek Parfit (1984) in p.312, footnote 2) that some psychological relation, but not identity, is what matters in survival. Somehow this leads her to defend the claim that only facts about the actual future existence of the foetus have normative significance. Thus, if it can be reasonably expected to become a person, we may have reasons against killing it. Contrariwise, if it cannot be expected to become one (for example, because the pregnant woman has decided to have an abortion), no such reasons exist. Since Harman believes that the only reasons against killing the foetus are those given by facts about its actual future life, she concludes that once a woman has decided to end her pregnancy "early abortion requires no moral justification whatsoever" (p.314). I find her view deeply problematic. First, suppose that she is right in believing that the only reasons against abortion that should factor in the woman's deliberation are those given by facts about the actual future of the foetus. Why should it matter at all, when deciding whether to abort an *early* foetus, whether it shall become a person? Harman accepts that what matters in survival is psychological continuity and the early foetus is, by Harman's definition, psychologically disconnected from its future. That is true whatever the woman decides to do with her pregnancy. Thus, what Harman ought to have claimed is that early abortion requires no justification whatsoever irrespective of whether the woman chooses to carry on with her pregnancy or to terminate it. In the second place, one could criticise her neglect of impersonal values. Later on in her article she defends the view that what does require moral justification is

What about the late foetus and the infant? It will depend on how demanding our account of the relations that ground prudential concern is. Let us grant that the weakest psychological relation justifies the attribution of some minimal time-relative value. From the standpoint of practical reason, it is not true that these cases of foeticide and infanticide are exactly like contraception, for in such cases there is actually some being which enters in the relations that matter with some future good. But even then, the foetus or infant is very weakly psychologically related to its future. It is useful here to remember Valjean-Madeleine's life in *Case Three*. In this scenario Madeleine's goods were for Valjean relevantly like in *Case Two*, where they were two numerically distinct individuals. Something similar obtains between the late foetus and its future –as McMahan stresses, it is much like somebody else's:

> "The developed fetus cannot envisage or contemplate its future and hence cannot have future-directed psychological states, such as intentions; it would, if it were to become a person, be unable to recall its life as a fetus; and it now has no psychological architecture—no beliefs, desires, or dispositions of character—to carry forward into the future. It is, in short, psychologically cut off or severed or isolated from itself in the future."[158]

---

to fail to have an abortion, since that creates a person with whom the mother shall have a "unique responsibility and relation" (p.324). We may safely assume that here Harman refers to a responsibility regarding the well-being of that future person. But, as the Non-Identity Problem shows (see CH. I, **§4**), there are situations in which we can only explain why creating a person is morally objectionable by appealing to impersonal values. If that is so, and values of this kind provide us with reasons to have an abortion, they should also factor in a woman's deliberation during the early part of her pregnancy by giving her reasons against terminating it when the person the foetus would become would expectably have a life worth living. But then it would be false that early abortion requires no moral justification whatsoever. However, that is, as I mentioned above, the position which Harman is intent on proving tenable.

[158] McMahan (2002, pp. 275-276).

As I explain in APPENDIX B, it is still a matter of debate whether or to what extent the unborn experiences some psychological development in the womb. If it does, then this account can vindicate the widely shared intuition that our reasons against foeticide (as given by the personal value of foetal life) grow stronger as the pregnancy advances to its resolution[159]. If, on the contrary, the foetus remains in a sleep-like state in utero, this is an intuition we would most likely have to discard. Certainly, the newborn infant is not, for its part, as weakly related to its future as the late foetus. Nevertheless, in the ways that matter foetuses and infants are related to their futures in a much more tenuous fashion than persons. Consequently, our reasons against killing them, as derived from the personal value of their lives, are correspondingly less weighty.

Denying that identity is what matters and that we should instead ground prudential concern on some psychological relation or other is, thus, a further way in which we can resist the view contrary to foeticide yielded by the *Future-of-value-Account*. I am aware that many people will find it very upsetting that on these views infanticide comes closer to contraception than to homicide. I shall deal with this problem, common to many ways of challenging a moral outlook contrary to foeticide, in this chapter's **§6**.

[159] Very early in the contemporary abortion debate, N. C. Gillespie (1977) argued that it is necessary for the truth of the belief that our reasons against foeticide (given by facts about the foetus itself) grow stronger as the pregnancy advances that there is some gradual property which the foetus possesses in increasing degrees as it develops into a paradigm instance of person. Alas, he failed to identify any such property. Later on, Paul Bassen (1982) envisioned the possibility that both identity and what he called *victimization* were gradual, just as mentality is. But it appears that he was sceptic about the possibility of finding a principled way of deciding whether that could be so. And yet, Warren Quinn (1984) proposed that things may enter gradually, instead of suddenly and entirely, under a substance sortal. Thus, it would be *strictly* true that a foetus is only *partially* one of us. Even though Quinn believed that identity matters, his ontology allowed him to discount the foetus's loss in dying to the extent it was not completely the future person who would live that future.

§4. *On claiming that a desire to live is necessary for someone's death to be prudentially bad.*

In this section and the next one I shall review several arguments for the normative significance of a desire to continue to live. In different ways these arguments attempt to establish that our most important reasons against killing individuals with an appropriate desire to live are of a different, more stringent sort than those against killing other individuals. On one view, this is because death can only be prudentially bad for someone in case she desires to live. On another view, this is because it is believed that what matters most in the morality of killing persons is whether the victim desired not to die, rather the extent to which her future life would be worth living. For now, though, I shall restrict myself to various versions of the first of these views –that which accounts for the importance of a desire to live in terms of the loss suffered in dying. Thus, to the premise in the *Future-of-Value Account* which states

(b) The badness of death for the one who dies at a certain time is determined by the net value her life would have were she not to die at that time.

These authors add

(f) Prudential value is desire-based.

For suppose that we accept that our most important reasons against killing a person are given by the fact that dying would be bad for her. Also let us accept that the badness of death for the one who dies involves the severing of the prudentially relevant relations between an individual and a worthy future. It is here claimed, further, that for there to be reasons against killing someone, as given by what is valuable for her, it is *necessary* that she desires to continue to live. On such a desire-based account of prudential values, whenever an individual has prudential reasons to try and make it that some state of affairs

obtains (e.g., the state of affairs of her continuing to live) it is either because its obtainment is the object of some telic desire of hers or because it is the object of some instrumental desire, that is, it is necessary for the obtainment of some further state of affairs which is the object of a telic desire. Thus, it can only be true that ceasing to exist is disvaluable for someone because it deprives her of some future life when continuing to exist is the object of a telic desire or when the contents of that future live are otherwise necessary for the satisfaction of a telic desire. If, additionally, we deny that foetuses possess the appropriate desires, then we may conclude that our reasons against killing them are of a different kind and of an inferior strength than our reasons against killing a person[160].

These kinds of desire-based accounts of the badness of death need to tell us, first, what properties a telic desire ought to have in order to be reason-giving[161]. If they are to be at all plausible, it must be admitted, for starters, that whether a particular desire is, at some time, *occurrent* or just being had *dispositionally* does not matter for prudential purposes. Someone's desire is occurrent at some time in case that individual is consciously entertaining it at that time. Thus, for example, when reflecting about the basic aims she wants to achieve in her life Clara thinks 'I want to be a respected neuroscientist'. But the mere fact that she stops thinking about that aim to reflect on what to cook for dinner, or falls asleep or unconscious, does not make it true that being a respected neuroscientist is something Clara no longer desires –it is simply the object of a dispositional mental state.

Moreover, it would seem that not all of someone's *actual* telic desires, those someone has in the actual world, necessarily matter for prudence. The desires

[160] The following is an important question: is it *true* that the obtainment of states of affairs is good for someone *because* that one desires their obtainment? A conditional reply would be: "if Subjectivism about reasons, in general, is true, then Subjectivism about prudential reasons, in particular, must be true". Yet is it true? I do not know, though I am inclined to Objectivism. On this distinction see CH. I, **§2**.

[161] For the following distinctions among desires I follow Boonin (2003, pp. 64-79).

to be considered reason-giving are rather those (merely) possible desires which an individual would form under *ideal* deliberative circumstances. Otherwise it would have to be admitted that *any* telic desire actually had by an individual gives her reasons for acting. But surely we can conceive of cases in which someone actually has a desire to die and yet we deny that such desire (or the telic desire to which it is related as an instrument) is reason-giving; we may even want to claim that, in spite of her actual desire, such individual has reasons to live. For suppose Clara was indoctrinated since childhood in the normative believe that her supreme rational aim was the good of some cult, or some political faction, and the situation came about that her death was truly necessary for the good of that cult or faction. Or suppose that it is some mental disorder what causes her to form the normative beliefs that make her death, under certain circumstances, instrumentally desirable for her. On a desire-based conception of reasons we cannot denounce her telic desires on grounds of having an inappropriate object, but we may denounce them on grounds of the inappropriateness of the procedure of which they are the result.

Can we attribute to an early, non-minded foetus *any* ideal telic desires? We can truly say, at least, that if it is not killed he *will* have some desires –that is, we can attribute to him possible future desires. Do these make it true that it presently has some instrumental desire to continue to live? I do not think so. When the actuality of some future telic desire is not dependent on a present choice of action I see little difficulty in admitting that it may extend its force backwardly to the past, generating present instrumental desires. For suppose that some enhanced rational agent had, at the time of choosing a course of action, knowledge not only of her present telic desires, but also of her future ones. Surely it would be irrational of her not to take those into account in her present deliberation. But this is exactly what we are imagining we can do in this case, namely, that we can envisage the future desires of a presently unconscious foetus. If tasked with acting for its sake it seems to me that future desires ought to enter into the practical calculus. But this simply follows from the more general claim that the actual desires of future people are reason-giving:

our concern ought to be with the desires existent at the time when the effects of our action occur, not at the time our action is performed. In this case, however, the actuality of the foetus's future telic desires is indeed *dependent* on the choice to be made. Suppose it is killed: what shall never be desired, in a telic fashion, by the foetus cannot make it instrumentally desirable for it to continue to live. And since nothing will be desired by it, it cannot be presently true that its life has any prudential value for it. It is worth noting that in this reasoning nothing was presupposed about the significance or lack thereof of numerical identity, or about what it might consist in. True enough, different answers to these questions entail different correct descriptions of our practical dilemma – whether to kill one of us as opposed to whether to prevent one of us from existing. Whatever the right answers to those questions are, the future desires of the non-minded foetus are, relative to our choice, non-actual.

The foetus, however, shall eventually acquire a mind –at a minimum, the capacities for experiencing pain and pleasure. Let us assume that concomitant with those the capacity exists of desiring that the pleasure be experienced, while it is experienced, and desiring that the pain be not experienced, while it is experienced. Whenever they are entertained, these desires are, of course, both occurrent and actual. They also seem to be of the sort to make pleasure and pain matter prudentially for the foetus. Suppose that the value of the pleasure which that organism shall enjoy throughout its existence will greatly outweigh the disvalue of the pain it shall suffer. Is it then justified to attribute to it an instrumental desire to continue to live?

I am afraid that now the problems of identity and its normative significance are consequential in order to find the right answer to that question. If we believe that we are essentially some living organism and that numerical identity is what matters, we can modify Marquis's argument so that it yields a conclusion contrary to foeticide only for the late, minded foetus. This seems to be David

Boonin's proposal[162]. After conceding that the infant (or the conscious foetus) cannot conceive of itself as a temporally extended entity that ceases to exist in dying, he adds that

> "[I]f he did understand these things, he would surely desire that his future personal life be preserved since he would understand that this is necessary in order for him to enjoy the experiences that he does already consciously desire to enjoy."[163]

As he notices, other beings possess fairly simple actual desires similar to those the foetus or infant has[164]. He denies, however, that this fact commits him to the claim that killing those beings is just as wrong, either in degree or in kind of wrongness, as killing a person[165]. It all depends on whether the future of the *particular being* in question is sufficiently like ours in value. Now, since here we are not arguing on an objective, value-based conception of reasons, in order to determine whether some future life is as sufficiently valuable as that of a person, we cannot merely inspect at its contents. We must recognise further

[162] Boonin (2003, pp. 56-85, 115-129). His adherence to Animalism can be inferred from his claim that the foetus captured in a sonogram image is "the same little boy [as his son] at a very early stage in his physical development" (p. xiv). Also, Boonin is, as he explicitly claims, just trying to present a more parsimonious and salient version (pp. 73-79) of Marquis's account and at no point he questions its assumptions on the nature and relevance of our identity. Incidentally, Boonin does not seem to be aware that he is misconstruing Marquis's argument as resting on a desire-based conception of prudential reasons. Thus, Boonin claims that Marquis believes that the wrongness of killing can be explained by appeal to a being's *present and future actual desires*, whereas he can explain it by appeal to *present ideal desires*. By dispensing with future desires, Boonin's account would be the more parsimonious one; by appealing to ideal rather than actual desires, it would be the more salient one. But as Marquis himself protests (2007a, pp. 210-212), his account was neither expressed in the language of desires nor needs it. I concur with Marquis that, being that the case, it is highly dubious that Boonin's alternative proposal is actually more parsimonious or salient than Marquis's.

[163] Boonin (2003, p. 84).

[164] Boonin (2003, p. 84 n. 36).

[165] Because killing a person violates a deontological constraint. I shall dwell on this aspect of Boonin's account after considering Michael Tooley's proposal.

that, given their psychological capacities, the range of events most animals are capable of experiencing is much limited than that of persons and that it is only within this range that the set of events whose occurrence (or non-occurrence) can be the object of a being's ideal telic desires is found. When assessing the prudential value some life has for the being whose life it is, then, we have to determine the extent to which its actual contents shall satisfy those telic desires –that is, to what degree its life is an instrument for the satisfaction of said desires. This reasoning is what allows us to conclude that, provided that the foetus or infant has a future life similar to ours, and assuming that it will develop psychological capacities similar to ours, we are justified in attributing it an instrumental desire to live much like ours. Therefore, our reasons against killing it, as derived from the prudential value of its life, are much like those we have against killing a person.

Thus, by grounding Marquis's *Future-of-Value Account* on a desire-based conception of personal reasons, the killing of the early, non-minded, foetus can be excluded from the scope of its conclusion contrary to foeticide. The scope of said conclusion can be further restricted, however, if we reject –on some of the grounds reviewed in the previous sections- the account of personal identity or of its significance which Marquis's theory, as well as Boonin's modified version, presuppose. For example, we saw how it can be argued that the extent to which some *future* desire or event can be said to *presently* belong to some individual is determined by the degree in which the relations of psychological continuity and connectedness hold between said individual now and herself at the time when she shall entertain the desire or partake of the event at issue. In those cases in which radical psychological discontinuity obtains, those future desires or events belong, from a prudential standpoint, in somebody else's life. As I argued in the previous section, however, numerical identity is not what matters. By committing this very same error, Boonin's modified *Future-of-Value Account* is unfit as an alternative to Marquis's original theory

Contrariwise, the intuition that identity is not what matters seems to underlie Michael Tooley's argument for the permissibility of foeticide and infanticide[166]. Tooley distinguishes between those beings who are mere momentary subjects of experiences and those who are temporally extended. For some subject of experiences to be considered as extended over some time it is not only necessary that its various temporal stages[167] be causally and psychologically continuous, but also that these be unified over time in ways only possible through the exercise of the concept of a continuing self[168]. It must be noted that this is distinct from the requirement that these stages be unified by the presence of the belief that they all belong to the same subject of experiences, which is the way self-consciousness is usually understood. For it is conceivable that a self-conscious entity lacks desires throughout its existence; also, the concept of a continuing substance can be exercised in ways other than by having that sophisticated belief[169]. Imagine that some entity not only desired that pleasure be experienced, while it is experienced, but that it also expected the pleasure to be experienced, even if only for a brief time, after the pleasant sensation has ceased. During that brief period, the various temporal stages of that entity would count as a single temporally extended subject of experiences. On the other end of the spectrum, persons such as your or me count as a single subject of experiences for a long stretch of time (typically, for all our lives) because our successive temporal stages are unified by a sophisticated net of mental states involving the concept of a continuing self –thus, we actually believe ourselves to be temporally extended and we harbour desires, hopes and expectations for the further future, as well as memories of past such mental states. Whenever the

[166] Here I refer to Tooley (1983; 2009), which represent a revision and expansion of his argument in Tooley (1972). It must be noted that in his more recent work he explicitly claims that persons are constituted by, and non-identical with, a human organism (2009, pp. 51-59), and less explicitly, but clearly enough, also assumes that stance in dealing with several objections (2009, pp. 21-23, 30-32, 40).

[167] By 'temporal stage' I merely refer to some period in the existence of a persistent entity, without thereby committing myself (or Tooley) to an ontology of temporal parts.

[168] Tooley (1983, pp. 123-146).

[169] Tooley (1983, pp. 133-134, 145).

continued existence of a person is necessary for the satisfaction of desires of hers obtaining at other times in her life, she ought rationally to form the instrumental desire of continuing to exist.

According to Tooley, if an individual ought to have an instrumental desire to continue to live, then ceasing to exist is prudentially bad for it and, hence, there are strong reasons against killing it. On the other hand, if the various temporal stages of some minded entity are radically disunified, it does not matter prudentially for any such entity-stage whether future entity-stages of the same entity shall obtain whatever is good for them and avoid whatever is bad for them. Indeed, for any such entity-stage to be prudentially concerned (if that were possible) for any future entity-stage of the same entity would be as absurd as it being prudentially concerned for some entity-stage of a different entity. One implication of this is that, caeteris paribus, we ought to be indifferent between an outcome in which some future good accrues at some time to *x* and an outcome in which some similar future good accrues at some time to *y*, where *x* and *y* are both mere momentary subjects of experiences. Suppose that *x* is an already existing entity and that, unless we intervene, the actual outcome will be *x*'s enjoying the good. Suppose further that the way of changing the outcome would consist in some action that causes the destruction of *x* and the creation of *y*, which would then be the one enjoying the good. The fact that *x* is destroyed gives us no reason against intervening, for it is false that in the sense that matters it has lost some future good. All that matters in this choice is that no good be lost and that, therefore, *someone* enjoys it. Our reasons for choosing would be, then, impersonal in character.

Tooley is well aware, however, that the unification over time of a subject of experiences admits of degrees. He is consequently inclined to think that this makes it possible that there exist subjects of experiences –those he calls *quasi-persons*– which, though temporally extended, are so in a more restricted way than persons[170]. It is plausible that some non-human animals are quasi-persons

[170] Tooley (1983, pp. 300-302, 407-412).

in this sense, even if they are temporally extended only towards the near future and the near past. Our reasons against killing these animals, whenever it is instrumentally desirable for them to continue to live, would be given by the fact that death is bad for them. Yet, it does not seem plausible to Tooley, on the evidence available, to affirm that a newborn infant, much less a foetus, qualifies as a quasi-person[171]. If they are to be considered, on Tooley's terminology, mere momentary subjects of experiences, then our reasons against killing them cannot be given by the fact that death would be bad for them.

Before proceeding to other desire-based accounts, it is convenient to consider here a further feature common to Tooley's and Boonin's view that I have hitherto not cared to mention and which would be unjust to omit –they are both *deontological* accounts of the wrongness of killing persons and both place great stock in the notion of a *right*, even though they have quite different conceptions of what a right is. For instance, Boonin wants to say that from the moment we gain consciousness as a foetus, we possess a right to life. According to Boonin[172], the notion of a right plays the following role in moral reasoning: to claim that an action violates a right is to say that it is morally impermissible and that, therefore, an agent has decisive reasons against choosing it, even it is optimific according to some theory of the good. Being morally impermissible is one way an action might be morally wrong. Another such way is for an action to be permissible, but criticisable. This means that whenever there are decisive reasons against choosing it, these are given by facts other than that it involves the violation of some right. The wrongness of morally impermissible actions, then, would not only be of a higher degree than the wrongness of merely criticisable ones, but also of a different, graver kind. Whereas there is a deontological constraint against killing beings with a future like ours, no such

[171] Tooley embarks upon a lengthy review (1983, pp. 357-407) of the scientific evidence available on human psychological and neurophysiological development, such as it was at that time, in order to conclude that newborn infants cannot even qualify as quasi-persons.

[172] Boonin (2003, pp. 3-9).

constraint exists against killing the preconscious foetus[173] and those conscious beings whose future is sufficiently dissimilar to ours[174].

As to Tooley, at first he uses the language of rights, thus speaking of a right to life. But later in the course of his argument he recognises that 'right' is not a basic moral notion[175] and analyses someone's having a right to life in terms of it being intrinsically wrong to kill that one because it is against its interests to die[176]. Here 'intrinsically wrong' means that there is a prima facie obligation not to kill it (or, as I would have it, that there are strong reasons against killing it). Thus, rights are not to be understood (à la Boonin) as providing us with decisive reasons against killing its bearer. Yet, in the latest iteration of his argument, Tooley considers himself a deontologist[177], and makes us privy to what he deems an entailment of such stance: considerations of value maximisation cannot justify killing those who possess "a serious right to life"[178]. Also, as we have seen, Tooley believes that there are various degrees of quasi-persons and correspondingly more or less serious ways of possessing a right to life. Perhaps these correspond to more or less stringent deontological restrictions against killing –howsoever this might operate in practice. He denies, however, that the foetus or newborn infant are protected by such restrictions, being as they are mere momentary subjects of experiences. According to Tooley, whether there

[173] Notice that the claim that conscious foetuses have a right to life is compatible with the claim that foeticide is in many cases, in Boonin's terms, morally wrong just in the sense of being merely criticisable, but not impermissible. One just needs to argue that in many cases, despite appearances to the contrary, foeticide by abortion does not constitute a violation of the foetus's right to life. This is exactly what Boonin (2003, pp. 133-281) tries to do by painstakingly defending Thomson's (1971) argument from many of the objections it has received over time.

[174] It is not clear to me whether, according to Boonin, merely conscious beings with a future sufficiently dissimilar to ours have *no* right to life or merely a right to life which grounds a *less stringent* deontological constraint against killing them.

[175] Tooley (1983, p. 95).

[176] From now on I merely follow Tooley (pp. 95-116).

[177] Tooley (2009, pp. 50-51).

[178] Tooley (2009, p. 41).

are reasons against killing them will depend on whether we have reasons to create additional happy people, but he is sceptic that such reasons exist[179].

I would like to briefly mention one final variation, Peter Singer's, of the view that explains the importance of a desire to live in terms of the loss suffered in dying. Unlike the positions reviewed thus far, it is markedly consequentialist –what we ought morally to do is to bring about the most valuable outcome. On this view, some values are desire-based, and are related to what is good or bad for an individual[180]. Thus, death can only be bad for an individual in case it frustrates some of her desires[181]. Of course, that is only possible if the being in question has desires regarding her future, so that only beings to some extent aware of their continued existence can suffer a loss in dying. Human persons are the clearest example of a being with many and important future-directed desires. Evidence suggests, though, that they are not the only ones, even if they possess a greater degree of self-awareness than other sentient beings[182]. The greater the capacity some being has for having desires about her future, the more desires that can be frustrated when it dies, and the stronger the reason we may have against killing it. It must be noted, however, that according to Singer there is nothing positively valuable in the *mere* satisfaction of desires – satisfied desires do not add value to the world. Desires matter morally because

[179] Tooley (1983, pp. 242-284). I shall be brief about this. On Tooley's population ethics the *quantity* of worthwhile lives never matters. Only two things matter, and *negatively* so: the quality of lives (lives not worth living) and inequality (unequal opportunity to acquire the means that make possible a satisfying life). Thus, though we have reasons against creating lives in which these things figure, that they shall be absent of some possible life does not give us reasons to make it actual. As we shall see below, Singer (2011) reasons along similar lines.

[180] Singer (2011, pp. 116, 119).

[181] Singer (2011, pp. 73-81, 100-104, 122). These are *ideal* desires: "The version of preference utilitarianism that I have tried to defend is based on the desires we would have, if fully informed and thinking calmly" (2011b).

[182] Singer (2011, pp. 94-100).

unsatisfied desires are objectively *disvaluable*[183]. This has some nice implications. Suppose that we determined that we could make sure that some future person shall have none or very few of her desires unsatisfied, but that in order to introduce her into the world it was necessary to kill another person, the life of which is worth living, but who would otherwise die with many more of her desires unsatisfied. None of the satisfied desires of the possible future person give us reason to create it, because that would not make the outcome better. And yet, the desires of the already existing person give us reasons not to kill her, because their dissatisfaction would make the outcome worse. This is the sense in which, in spite of our duty to maximise desire-satisfaction, persons are not replaceable.

We may also have reasons not to kill merely conscious beings, but most of the time these are given by impersonal values not based on desires[184]. Suppose that pleasure is objectively valuable, whereas pain is objectively disvaluable. Merely conscious beings serve as vehicles for the introduction of pleasure and pain into the world. Though it would be irrational to preserve their lives for their own sake, we still ought to bring about the most valuable outcome. In order to achieve that aim, sometimes we will have reasons to bring conscious beings into existence, or to preserve the life of some conscious being. Conversely, some other times we will have reasons to prevent the existence of some conscious being or to kill it, even if in the way in which it shall suffer the least. But, as I pointed out when discussing a similar point in Tooley's theory, we ought to be indifferent as to which merely conscious being has a certain pleasurable experience, in so far as that does not affect the overall value of the outcome. But suppose that we determined that we could increase pleasure in the world by introducing some conscious being in it and that in order to do so it

[183] This is what Singer calls the "debit view of preferences", which is, as it will become apparent, very closely related to his claim that merely conscious beings are "replaceable" whereas beings who are self-aware are not (2011, pp. 104-114).

[184] Singer (2011, pp. 85-90), but see also the reference to the discussion of replaceability in the previous footnote.

would be necessary to kill another conscious being, the life of which is worth living. That latter fact would give us no reasons against killing it. Supposing that there are no other reason-giving facts to consider, we would have decisive reasons to kill that conscious being in order to introduce the new one. It is in this sense that Singer claims that merely conscious beings are replaceable.

What to make of foeticide on this view? Since foetuses cannot have desires about their future, it cannot be said that death is bad for them, or that they lose some future life, however worth living, when they are killed[185]. What reasons we may have will be given, as it is also the case regarding possible future beings or merely conscious ones, by the extent to which its life shall serve as a vehicle for introducing valuable things into the world –that is, our reasons are impersonal in character[186]. Hence, provided that its life will be worth living, we have some reasons against killing it[187]. It is implausible, however, that they will be decisive in most cases, or as strong as some view contrary to foeticide would have them to be. For, first, one should assume what is doubtful, namely, that in our present circumstances we *all* have decisive reasons to add further people into the world. Second, one ought to argue that *each* woman has decisive reasons to bear some number of children. But even if women had such decisive reasons, surely it would not matter *when* they bear their children, provided that that does not have a negative impact on the children's well-being. As Singer puts it

> "This argument does not provide any reason for thinking abortion worse than any other means of population control. If the world is already

[185] Singer (2011, pp. 134-138).

[186] Singer (2011, pp. 139-140).

[187] What if the foetus is conscious and the method of killing can cause it pain? That is bad, so we have reasons to choose a painless killing, if possible. If it is not possible, then we have additional reasons not kill the foetus. But Singer (1995, pp. 209-210) believes that even in these cases the reasons in support of a woman's choice can be strong enough to make the killing a late foetus permissible. This is, of course, assuming that foetuses can experience pain in utero, which is disputed, as I explain in APPENDIX B.

> overpopulated, the argument provides no reason at all against abortion."[188]

It must be noted, then, that if we adopt a desire-based conception of personal reasons, there are several ways to deny that killing a foetus is bad for it. In so doing we can restrict the scope of an account of the wrongness of killing such as Marquis's, which claims that our main reasons against killing are given by facts about the prudential badness of death. If we merely shift from a value-based conception of prudential reasons to a desire-based one leaving the rest of the argument's premises intact –as, for instance, Boonin does- our conclusion can only be that death cannot be bad for the preconscious foetus. Yet it would be a mistake to proceed in that way, for identity is not what matters. Those positions which refuse to concede normative significance to identity may extend that conclusion to the conscious foetus and the infant. As Tooley and Singer admit, that does not preclude the existence of impersonal reasons against killing foetuses whose lives will be worth living. But even we have such reasons, it is doubtful that *in fact* in most cases we ought, all things considered, to refrain from foeticide.

As I advanced, though, there is another way to attach importance to a desire to live, one which does not consist in making such a desire necessary for the badness of death. It has been claimed that what matters most in the morality of killing is whether the victim desired not to die, rather than the extent to which her future life would be worth living. I turn my attention to this view in the next section.

§5. *On denying that killing us is seriously wrong merely because it causes a prudentially bad death.*

[188] Singer (2011, p. 140).

In this section I shall comment on a last way to object to the *Future-of-Value Account*. In some important sense, this is one of the most radical rejections of that account, since it consists in denying

(a) The fact that dying would be bad for the one who dies gives us our main reasons not to kill her.

Most writers I have discussed accept the truth of (a). They merely dispute what else must obtain for someone's death to be as bad as we believe the death of a person is for her. In rejecting (a), though, we are instead declaring that in those cases in which we believe we have our strongest reasons against killing it is false that these are given by the prudential value of the victim's future life. If we can show that foetuses lack whatever provides us with these strong reasons against killing, we will have come a long way in arguing that our reasons against foeticide are weaker than the views I am criticising would claim.

What is the motivation for severing the link between the badness of death and the wrongness of killing? Most people have a strong intuition that whenever killing is wrong, it must be because it is bad for the victim. From here there is a short step to the claim that it must be wrong because *dying* is bad for the victim. After all, causing death is the most conspicuous effect of killing. But this easily lends itself to contradiction with a further belief held by most people, namely, that whenever killing a person is wrong, it is equally wrong to kill *any* person, all else being the same –where 'all else' does *not* include the prudential value of the person's future. Sadly, if our most important reasons against killing are given by the badness of death and the badness of death correlates with the prudential value of a person's life, then that cannot be true. On the contrary, our reasons against killing someone with a life of moderate value will be weaker than those against killing someone with a life of great value. All things being

equal, and confronted with a choice, it would be less wrong to choose to kill the former rather than the latter[189].

If that is the conclusion we want to avoid, then there ought to be a property possessed by all persons such that (a) does not vary in degree or (b) whose gradual variation somehow does not translate into differing strengths in our reasons against killing. A favoured starting position is the following. There are, as we have seen, different ways to argue that a life is preferable to death for some creature, as well as different ways of saying that a creature has a desire to live rather than die. Persons stand apart, though, in that, unlike other creatures, they may also *choose* to live, or not to die. They certainly have the capacity to so choose. At least in most occasions in which they actually choose to live or, perhaps, whenever such choice would follow from the system of their other normative beliefs and desires, we have strong reasons to honour their decision. We can safely assume then that killing them is a particularly serious manner of failing to honour their choice not to die.

Now, a person's choice is typically the result of the exercise of her distinctive faculties for practical reasoning. Even if most non-persons respond to practical reasons, persons possess the additional capacities of forming beliefs about what the relevant reason-giving facts are, the strength of the reasons with which these provide her and what she has most reason to do. Persons are, in a word, rational. How is it valuable, that is, reason-giving, that a being be rational? One way to argue for this follows a tradition initiated by Kant[190]. It seems that Kant distinguished between two kinds of values[191]. Some events are valuable in the sense that they are *good* –either, as we would say, personally, or impersonally or both. We are familiar with goodness. Whenever an event is valuable in this sense, we have reason to *promote* it. Thus, what we have most reason to do,

[189] At least this is the motivation that compelled Jeff McMahan to seek an alternative account of the wrongness of killing persons.

[190] As recognised by McMahan (2002, p. 252) or Singer (2011, pp. 83-84).

[191] Here I follow Parfit's (2011a, pp. 233-244) reading of Kant, and use some of his examples..

as per this kind of value, is to bring about the best possible outcome –the one containing those events the composition of whose value yields the highest result.

Besides goodness, Kant identified a different sort of value things may have, which he called 'worth' or 'dignity'. Whenever some existing thing is valuable in this way, we have reasons to *respect* it. What is required of us when something is valuable in this way depends on the kind of thing that is the object of respect. Sometimes the appropriate way to respect something will consist in trying to ensure its continued existence. This might be true of the majestic, old tree Clara was so impressed with or the works of art her father Peter admired. But that is not always so. It is oft said that the corpses of persons merit respect. Assuming that it is so, an appropriate manner of acknowledging it would consist not in preserving them, but rather in, for instance, "burning them bedecked with flowers on some funeral pyre, rather than throwing them onto some rubbish dump" [192].

A possible way of developing Kant's thought, embraced by Jeff McMahan[193], would be to claim that persons are a proper object of respect –i.e., have worth or dignity– because they are rational[194]. Now, rationality consists in or presupposes a number of higher psychological capacities, such as self-

[192] Parfit (2011a, p. 238).

[193] I will here present the argument in McMahan (2002, pp. 240-265, 473-485; 2008, pp. 93-104).

[194] There seems to be some disagreement about what Kant actually believed. Parfit (2011a, pp. 241-242) is of the opinion that Kant indeed believed that *persons* have the value that demands respect, though he also believed that our capacity of acting morally and having a good will was infinitely *good*. Yet McMahan (2002, pp. 248-249) claims instead that Kant believed that *rationality itself* has the value that demands respect. This view reveals itself as inimical to his and most people's intuitions, for it forces upon us the conclusion that we have decisive reasons against rational suicide and euthanasia (except, perhaps, to prevent the degradation of our rational faculties through dementia). This is why he sees himself as providing a more plausible version of Kantianism, rather than as being faithful to Kant.

consciousness, autonomy, analytical intelligence and imagination[195]. McMahan suggests that there is some point along the scale that measures psychological capacities –what he calls the *threshold of respect*- that separates conscious beings into two spheres of morality, governed each by a different set of principles. Our treatment of beings that lie below the threshold is governed solely by reasons given by facts about what is good for them –this is the sphere of the *morality of interests*. On the other hand beings that stand above the threshold have the kind of value that demands respect. Regarding these beings, we are not only required to have an appropriate response to what is prudentially good for them, but also to the choices these beings autonomously make, or at least their choices about the most important aspects of their existence[196].

If my construal of McMahan's mind is correct, it would be a mistake to believe that the morality of respect is simply the morality of interests with a further layer of concern for autonomous choices. For one, "[m]orally, the gap between those above the threshold and those below is immense"[197]. Restricting his reflection to our discussion, this means that our reasons against killing a being that merits respect are significantly more serious that those against killing beings below the threshold. This is explained because of the "entirely different set of considerations" that governs this moral sphere, which makes it the case that maximisation of overall outcome value is no longer our sole moral aim. Thus,

[195] McMahan (2002, pp. 242-243, 256, 261). What are precisely the bases or grounds of the respect which persons merit? McMahan remains agnostic on this, but believes that the following view, which he attributes to Warren Quinn (1984), is true: "that respect for a person is closely connected with respect for the autonomous determinations of that person's will; therefore, autonomy must be a significant element of the basis of the worth that demands respect" (p. 260).

[196] McMahan (2002, pp. 245-246, 257-258). In the sphere of respect, what is good for a being must matter as well. Otherwise, we could hardly account for our intuition that not all acts contrary to what a being has autonomously chosen are equally wrong –e.g., tweaking someone's nose versus killing her. The strength of her reasons for the choice, as given by what is prudentially good for her or for those she is especially related with, ought also to be considered (2002, p. 257).

[197] McMahan (2002, p. 261).

the interests of the beings *below* the threshold "may be treated in a consequentialist manner", i.e., "traded against one another without any weighting other than for the strength of the interest itself"[198]. The beings *above* the threshold of respect, however, enjoy *maximum inviolability* –it is impermissible to sacrifice any of them with the aim of bringing about the impersonally best outcome[199], except in extreme cases[200]. Beings solely under the morality of interests lack this protection. Another consideration governing this moral sphere might be the *mode of agency*[201] by which someone's autonomous desire to live is frustrated: perhaps respect for that desire requires us not to kill, but allows us to let the being die. Whether an individual is below or above the threshold would be, then, most important for practical reason.

One problem, however, is that the high psychological capacities that are the basis of respect for persons vary among people –some possess them in a greater degree than others. Being that so, it would be natural to claim that the greater a person's capacity for forming true normative beliefs and acting on reasons is, the more she merits respect and the stronger the reasons to honour her autonomous choices, especially her choice not to die. Nevertheless, if that were so, our attempt to establish secure grounds for the belief that it is equally wrong to kill *any* person, all other things being the same, would have failed. McMahan suggests that one way to solve this problem is to consider the attributes that are the basis of respect as 'range properties'. Though McMahan does not provide us with a definition of the concept this expression stands for, he does points us to John Rawls's illustration of it:

---

[198] McMahan (2008, p. 98).

[199] McMahan (2002, p. 263).

[200] For example, McMahan (2002, p. 247) considers it impermissible "to kill one person as a means of preventing the killing of a greater number of others", though "if the number of others becomes sufficiently high, killing the one may become permissible". And he is clear that this is not a matter of balancing one autonomous choice not to die against a number of others, but against what is good for a number of other persons: "[i]n extreme cases, the demands of respect may be overwhelmed by countervailing pressures from the morality of interests". See also McMahan (2008, p. 98).

[201] McMahan (2002, p. 259).

> "For example, the property of being in the interior of the unit circle is a range property of points in the plane. All points inside this circle have this property although their coordinates vary within a certain range. And they equally have this property, since no point interior to a circle is more or less interior to it than any other interior point."[202]

McMahan suggests that autonomy could be regarded as a range property, at least when what the person chooses is whether to live:

> "If a person is sufficiently autonomous to know that being killed is incompatible with the plans he has for his own life, this is all that is necessary for his will to be autonomously set against his being killed. He is, where this matter is concerned, fully autonomous." [203]

In this way, whenever a person's choice against dying is fully autonomous we may have especially strong reasons against killing it. Moreover, someone's choice cannot be more or less *fully* autonomous depending on the degree of the psychological capacities a person's autonomy consists in. Thus, the strength of our reasons to honour fully autonomous choices does not vary across persons. All beings who have reached this level of full autonomy merit equal respect regarding their choices about continuing to live –they stand beyond what McMahan calls the *threshold of equal worth*[204].

It seems plausible to McMahan that the level of psychological capacities that situates a being in the moral sphere governed by respect is the same that elevates it beyond the threshold of equal worth. Let us follow him in calling, for the purposes of this section, those beings who reach at least such level of

[202] Rawls (1999 [1971], p. 444).

[203] McMahan (2002, p. 257).

[204] McMahan (2002, p. 249).

psychological capacities 'persons'. Perhaps I can now more precisely formulate McMahan's account of the wrongness of killing persons:

Intrinsic Worth Account:

(i) Proper acknowledgment of the respect which persons merit includes treating them in accord with the reasons given by what is prudentially good for them *and* the reasons given by their autonomous choices whether to continue to live.

(ii) These facts give us our main reasons against killing persons[205].

The clearest implication of this account is that whenever a person has autonomously chosen to live and living would be prudentially good for her we have significantly serious reasons against killing her. Another clear implication is that when a person's good lies in continuing to live but she has chosen instead, in a non-autonomous way, to die[206], proper acknowledgment of the respect she merits consists in disregarding that choice. Of course, we need an account of what it is for a choice to be autonomous, and anyhow there shall be cases in which that will be indeterminate. But the real problem cases arise when a person autonomously chooses not to live even though it would be bad for her

[205] It ought to be remarked, however, that McMahan (2002, pp. 264-265) does not conceive of the threshold that separates persons from other beings as "a sharp line with clear cases on either side", believing instead that "we all begin life lacking the kind of worth that commands respect, there is no point at which we instantaneously acquire it". There is an early period, whose end McMahan places somewhat beyond the first or second year of age, when we are "indisputably nonautonomous". There is a later period, starting around our sixth or seventh year, when we are "recognizably autonomous". During the middle period, when we are becoming autonomous, our status is indeterminate. He contemplates the possibility that some animals, such as higher primates, also possess this indeterminate status. McMahan suggests that, though these individuals merit fewer respect, in these cases the *Intrinsic Worth Account* applies too. Regarding these beings of intermediate moral status "the strength of [the moral] constraints varies with the level of psychological capacity of the individual to whom they apply" (2008, p. 98), for they lie below the threshold of equal worth.

[206] Perhaps because of "an aberrant mental condition" (McMahan, 2002, p. 259).

to die. Beyond strong intuitions regarding certain situations, McMahan seems unable to provide a full theory about this range of cases[207].

Yet, if his views on the clear case in which a person whose life is worth living chooses not to die are true, here we have a further way to argue that our reasons against foeticide –and infanticide- are much weaker than those against killing persons[208]. Foetuses and infants lie entirely in the morality of interests. This means, first, that the reasons against killing a foetus with a valuable future, as given by facts about the prudential value of that future, can never be as strong as the reasons against killing a person with a similarly valuable life who chooses to live, as given by the respect she merits. Second, these killings are governed exclusively by consequentialist principles, without the restrictions imposed by respect. But the reader may remember that on the *Time-Relative Interest Account* of the wrongness of killing, which is the one that McMahan applies to conscious beings outside the sphere of respect, our reasons against foeticide or infanticide were already very weak. The relations that prudentially matter hold to such a diminished degree between the foetus or infant and its future life that, from the standpoint of practical reason, it is *as if* that life was someone else's[209]. When McMahan's two-tiered account of the wrongness of killing is considered in its totality[210], the relative strength of our reasons not to kill a foetus or infant for its sake reveal themselves to be feeble indeed.

---

[207] Thus, he believes that one ought to prevent someone from committing suicide in some cases in which the person's choice to die rests on false beliefs, either normative (a mistaken view on the value of life; *ibid.*, p. 259) or factual (that extraterrestrials will carry her off to a better life; p. 479). And yet he believes that one ought to respect a prisoner of war's choice to face inevitable torture and death, rather than let his comrades painlessly kill him, "because he believes that it would be ignoble to evade torture in a way that would affirm the enemy's power over him" (p. 479).

[208] McMahan (2002, p. 339).

[209] McMahan (2002, p. 493).

[210] Or, alternatively, when one considers his three-tiered account by including in the theory those beings with "intermediate status" (McMahan, 2008, pp. 93-104).

§6. *On the permissibility of infanticide as a decisive objection to the permissibility of foeticide*.

For some of the theories that I have discussed in this chapter, it is strictly true that killing a foetus cannot be bad for it. For other such theories, even though they deny that this is strictly true, the foetus suffers so slight a harm in being killed that it gives us reasons against ending its life whose strength is similar to that of the reasons we would have if the badness of its death were impersonal. Now, it is troublesome for many people who would accept these conclusions that they also apply to the killing of infants. Some argue that this implication is reason enough to reject any theory from which it is derived[211]. What these people would like to believe is that our reasons against infanticide are as weighty as our reasons against killing persons. Perhaps they would even insist that these reasons be given by the *same* facts that make the killing of persons so seriously wrong in most occasions. Perhaps it can be shown that this discomfort offers no decisive grounds for rejecting any view favourable both to foeticide and to infanticide[212].

As Michael Tooley pointed out, if an appeal to a moral intuition is to be considered *decisive* in a discussion in moral philosophy, it must be shown that it

[211] This is, for example, Robert P. George's (2013) position, who regards the advocacy of the moral permissibility of infanticide as "madness". I believe that the following paragraph of his is worth quoting in full, for it expresses a widespread, popular opinion: "I really do believe that advocating the moral permissibility of killing infant children is scandalous. Anyone should immediately be able to see that killing infants because they are unwanted is unacceptable—even if they have trouble seeing, and are in need of information or argument to see, that killing a human being in the womb is wrong. Killing babies, like buying, owning and selling slaves (even if we debate whether certain labour practices are exploitative in ways that make them the moral equivalent of slavery), is not something that we should treat as worthy of being considered as a morally legitimate option".

[212] For the remainder of this section I will be chiefly following McMahan (2002, pp. 338-362, 450-455, 485-493), Tooley (1972; 1983, pp. 309-416) and Singer (2011, pp. 159-167), who have advanced the most robust defences of this conception of the moral status of infants.

possesses the epistemic value necessary to perform the argumentative role it is supposed to have[213]. After all, some intuitions, though not all, are proved false, or at least abandoned over time. Tooley suggests that a reliable moral intuition ought to have at least the following attributes. First, it must enjoy widespread acceptance, both diachronically and synchronically. Second, it must have been accepted because it appeared to be epistemically justified *in itself*. This excludes moral beliefs accepted merely because of prudential reasons or merely because they were an essential part of some comprehensive worldview which appeared to be epistemically justified. Finally, it must either consist in some *basic* normative belief –of the sort that claims that a certain fact is reason-giving–, or be derivable from other basic normative beliefs the person has which can be shown to be epistemically sound.

It seems to me that the belief that the moral status of infants is the same as that of persons lacks all these three attributes. Let us begin with its alleged widespread acceptance. Certainly that is true neither of all past societies, nor of all contemporary ones[214]. Historian John M. Riddle, when discussing infanticide as a method of population control, produces the following general statement:

> "Ancient law protected neither the fetus nor the newborn infant until there was acceptance by the parents, often by some ritualized tribal or

[213] Tooley (1983, pp. 24-30).

[214] Since others have done it, I see no value in reproducing here particular examples of such practices. But see Tooley (1983, pp. 309-322), Kuhse & Singer (1985, pp. 98-117), Singer (1995, pp. 214-215; 2011, pp. 153-154) or McMahan (2002, pp. 340-342). It is interesting to note, as these authors do, that even in contemporary, traditionally Christian societies, medical practice and parental preferences do not reflect the belief that the moral status of infants is the same as that of young children or that all infants have the same moral status. For example, sometimes parents refuse –a quite simple form of– surgery for their Down syndrome infant in order to remove a blockage in its digestive system (a fatal condition usually associated with that syndrome). Most probably they would not have refused it if they had born a normal child, nor if the condition had affected their Down syndrome child in its second or third year of age.

> community registration. In contrast, during the Middle Ages, religious and sometimes secular law spoke against the exposure of infants." [215]

Indeed, later on we are told that "the Middle Ages definitely deplored infanticide"[216]. This change in moral outlook is not surprising, since by that period Christianity had become the dominant religion in Western societies. Generations upon generations of Europeans, and later its cultural descendants in other continents, were educated in the Christian conception of the sanctity of life, opposed both to foeticide and to infanticide. Since it is but recently that the truth of Christian ethics has been questioned, it is not surprising that many people harbour an instinctive rejection of infanticide. This, of course, ties in with the second requirement, namely, that it can be confidently said that this belief gained acceptance because of its intrinsic epistemic appeal. The way in which moral beliefs about the killing of infants changed makes them suspect –in so far as we have reason to doubt the veracity of the religious worldview in which they are grounded, we have reason to doubt their epistemic value in deciding the issue of infanticide.

Finally, a strong intuition against infanticide cannot count as a basic normative belief. Someone who has that intuition could probably explain it in terms of more fundamental moral beliefs. For example, by identifying the property that foetuses, infants and adult human beings have in common and which gives us such strong reasons against killing them. That is, they must produce a general account of the wrongness of killing beings like us and submit it to discussion. When we consider all these features of our intuitions against infanticide, it is not unreasonable to suppose that it is unapt to function as a methodological touchstone. Invoking it cannot serve to settle the dispute –except, of course, if one cannot but accept a worldview of which it is a necessary part. In all other

[215] Riddle (1992, pp. 10-11). "Exposure of infants" is a translation of the ancients' way of referring to infanticide.

[216] Riddle (1992, p. 12), but see also his (1997, p. 18).

cases, though, its soundness must be established through philosophical argument.

A first point to note is one about consistency among our moral beliefs. Consider what happens when one reflects on the morality of foeticide in conjunction with that of infanticide and the killing of non-human animals[217]. If one adopts a view favourable to foeticide at any point in pregnancy, it is inevitable that in due course of reasoning the traditional moral beliefs in one or more of these other topics shall have to be revised. Suppose that one refuses to modify the beliefs that our reasons against killing infants, as given by the prudential value of their lives, are as strong as those against killing persons, whereas those against killing non-human animals, as given by the same facts, are much weaker (or non-existent). A first alternative is to appeal to species membership in order to account for our beliefs about the moral status of non-human animals. That will not do, however. For one, the foetus is also human, so that one would have to embrace a view contrary to foeticide. Also, we have already seen how it is implausible that species membership is morally relevant. One may, alternatively, adopt the view that identity suffices to ground prudential concern, and that we are living organisms. Since the psychological capacities most human infants develop grant them access to many goods which non-human animals cannot obtain, an untimely death is much worse for the former than for the latter. But then, again, one cannot adopt a view contrary to foeticide.

Suppose that instead of appealing to the future psychological capacities of infants, one appeals to their present ones. The first thing to note is that many non-human animals possess higher psychological capacities than those of non-premature newborn infants. If one wants to claim that these capacities suffice to make infanticide at least as seriously wrong as killing persons, one will have to admit the same regarding the killing of many non-human animals. Certainly, our reasons against killing animals are stronger than it has been hitherto admitted.

[217] This is the main point in McMahan (2013), but it was also noted by Tooley (1983, p. 423) and Mary Anne Warren (1984, p. 119).

To claim, however, that they are as strong as our reasons against killing persons is to commit oneself to the view that being a self-conscious, autonomous being capable of enjoying distinctively high goods has none, or very little, rational significance. I find that implausible.

The second thing to note regarding this strategy is that there is no such thing as *the* psychological capacities of newborn infants –what those capacities are depends on the stage of development of the foetus at the time when it was born. Some foetuses are born prematurely, and some are not. And, of course, not all premature foetuses are born at the same point in pregnancy. Given that different foetuses are born with different psychological capacities, how does one identify which are the relevant ones? An answer would be: those at birth, whatever they are. This seems, however, an arbitrary criterion. Birth in itself is a mere change in location, thereby being an implausible candidate for a property with such momentous significance. For consider what this would entail. Imagine two foetuses qualitatively identical in their psychological capacities, yet one remaining unborn and the other being a newborn. This theory implies that the same psychological capacities would give us serious reasons against killing the latter, but not the former. It appears to me that mere change in location cannot matter in this way.

Perhaps viability can help us to identify what those psychological capacities are. Viability refers to the foetal capacity to survive outside the womb, with or without artificial assistance[218]. It could be claimed then that the psychological capacities foetuses have at the time they begin to be viable are those that matter. The problem, again, is that this seems to commit us to attaching normative significance to properties which have none. Viability is relative to the technology available at the time and birth of place. Whether it is a good criterion for policy-making or not, this makes it dubious that it has any moral importance. For imagine again our two foetuses with identical psychological capacities, now

[218] There is a wide consensus that viability has in itself no normative significance. See, for instance, Boonin (2003, pp. 129-132), Singer (2011, pp. 126-128).

both remaining unborn. Now we are forced to claim that the spatiotemporal location of each of them may change the strength of the reasons against killing it given by its psychological capacities. Again, it seems to me that changes *merely* in space or time cannot matter for practical reason, much less change the significance which *other* properties have.

It must be considered now in what circumstances, if we are to have consistent moral beliefs, it ought to be considered that we have sufficient or even decisive reasons to kill an infant. I have presented three kinds of accounts of the wrongness of killing merely conscious beings entailing a view favourable to foeticide which are also favourable to infanticide. First, there were those theories which accepted that identity is what matters for prudential concern, but claimed that infants are merely our predecessors, not numerically identical with us. This might be held by one who embraced some version of the Lockean views presented in this chapter's **§2**. Then, in **§3**, it was proposed that we ought to accept that theory which argues from the weakness of the time-relative value of the infant's future life. Finally, there were those desire-based theories which argued that infants lack a desire to continue to live, such as we saw in **§4**. Since the facts considered relevant by these theories vary little or not at all across infants, then the proper way to distinguish between kinds of cases is according to the overall expectable value of the infant's (or future person's) life –no value, positive value and negative value. On the predecessor-view of infants and desire-based accounts, the overall expected value of the infant's life is wont to be considered impersonal in character, for the infant cannot be attributed a desire for its continued existence. On a time-relative value account, however, it must be admitted that one ought to attribute to the infant some prudential concern for its future.

Let us begin with those cases in which the infant's life will be devoid of all value – as is perhaps the case of infants born with anencephaly[219]. As I mentioned

[219] See McMahan (2002, pp. 450-455). But Steinbock (2011, p. 28) claims, after reviewing the relevant literature, that we cannot be certain that anencephalic infants are non-sentient. If they

several sections ago, anencephaly is the condition which consists in the absence of cerebral hemispheres. Since, presumably, our hemispheres contain the structures necessary for the realisation of mental states, this means that, probably, anencephalic infants are beings with no minds. From the standpoint of practical reason, they are like any other non-minded thing (such as rocks or plants) in that nothing can be good or bad for them. There are, then, neither reasons against killing the anencephalic infant for its sake, nor reasons provided by the impersonal value of its life, for it has none. Nevertheless, there might be compelling reasons against killing it. The organs of anencephalic infants can be used for transplants and, thus, to prevent the deaths of other infants whose life would otherwise be worth living. Even though it is true that, if provided with no artificial assistance, an anencephalic infant will live for only a short period of time, it seems that the way its condition causes its death damages its organs in such a way as to make them unsuitable for transplant. It would be better, then, to advance its death. Consequently, our reasons to kill the anencephalic infant for the purposes of harvesting its organs and saving other lives are weighty indeed. Provided that the infant's parents consent to it, I can think of no serious reason to the contrary[220].

Next we have those cases in which the infant's life is worth ending –that is, infanticide as a species of non-voluntary euthanasia[221]. Suppose, first, that we

---

were sentient they ought to be treated as infants with lives worth living, or worth ending, depending on the circumstances.

[220] As Singer (1995, p. 212) underscores, the preferences of the infant's parents in these cases ought to be considered of great import "because of the effects, both good and bad, that the continued life of their child will have on them and any other children they may have". See also Kuhse & Singer (1985, pp. 183-189). From this fact he derives the general rule that the decisions about the life and death of this infants should be made by its parents in consultation with the doctor, and not by "a judge" (by which we ought to understand 'a public official or organism').

[221] Non-voluntary euthanasia is that performed on a being which is "not capable of understanding the choice between life and death". This is distinct from involuntary euthanasia, in which "the person killed is capable of consenting to her own death but does not do so, either

have adopted an account that tells us that the reasons given by facts about the infant's future life are impersonal. Since, caeteris paribus, the possible world in which that infant exists is worse than that in which it does not, we have reasons to actualise the latter instead of the former. If we consider, in addition, the reasons given by the burden which such an infant may impose on its parents or on those who will have to take care of it, the case for infanticide in these situations appears to be quite strong. That is especially true if the cost of attending to the infant's needs shall preclude the possibility that its parents have another child with a life worth living.

Suppose now that we have adopted an account that derives our reasons for killing infants whose lives are worth ending from the value of those lives relative to the time we are to choose whether to kill them. It could be argued that an infant's psychological unity over time is so weak that even if the condition that would make its life worth ending –by, say, causing terrible pain- would obtain in a few days or weeks, our present reasons to kill it would not be significantly strong[222]. No matter how intense the future time-relative value of those experiences, it must suffer a severe discount because of the psychological detachment between the infant presently and those future events. This, of course, does not mean that we have no reasons to kill the infant for its sake. It means, nonetheless, that these reasons are quite weak. That is not what one would expect: intuitively, the intense, irremediable, future suffering of beings psychologically distanced from their future selves appears to give us strong reasons for killing them. A strategy I find attractive is supplementing these weak personal reasons with impersonal ones. Thus, it would be by combining these two kinds of reasons that we could explain, on this account, our sense of the badness of letting these infants live[223].

---

because she is not asked or because she is asked and chooses to go on living" (Singer, 2011, p. 158).

[222] This is a problem encountered by McMahan (2002, pp. 485-493).

[223] Interestingly, though McMahan believes that it "simply must be right" (2002, p. 493) that we have strong reasons against creating beings, or allowing the continued existence of animals, whose lives would be worth ending he is at a loss on how to adequately ground such intuition.

Last of all we find the cases of those infants whose life shall be worth living – may we have at least sufficient reasons in some of these cases to kill the child?[224] Arguably some instances of this kind of infanticide may be permissible, even if cannot be as easily justified as foeticide. The first point to consider is that the fact that there were sufficient reasons to kill a particular foetus doesn't mean that it will be justified to kill it after it has been born. Admittedly, the reasons given by the personal or impersonal value of its life shall remain more or less the same. As to the latter, it cannot have changed because of birth. As to the former, the time-relative value of the infant's life cannot have increased but slightly[225]. The circumstances of our decision, however, might have changed in a morally significant way regarding how other people's lives are affected by the infant's existence. Consider the burdens that an unwanted pregnancy imposes on the woman heavy with child. These are costs of which she can only relieve herself by terminating the pregnancy. If the foetus is not yet viable, abortion shall entail its being killed. If it is viable, it might be extracted from the woman's womb without killing it, but at a possibly greater cost for the woman because of the greater risks associated with this kind of intervention. Anyhow, on these accounts, the reasons given by the value of the foetus's life are so weak that on most occasions they shall be counterbalanced by the interests of the pregnant woman and other affected parties.

However, by the time that self-same foetus is born, it will no longer be true that the costs its existence might impose on its mother can only be relieved by killing it. Indeed, as it has been often pointed out in the literature, after birth there might be available other ways of distributing that cost[226]. In these cases, the

---

Partly, this is because, unlike me, he believes that there cannot be impersonal reasons to create beings with worthwhile lives, even if they are normally outweighed (2002, p. 492).

[224] See McMahan (2002, pp. 338-362) and Singer (2011, pp. 162-167).

[225] McMahan (2002, p. 342; 2013, p. 280).

[226] "This may be the principal reason why infanticide is, in general, more objectionable than abortion." (McMahan, 2002, p. 344); see also McMahan (2013, p. 280). Singer (2011, p. 154) seems to be of a similar opinion: "infanticide can only be equated with abortion when those

reasons for terminating the infant's life either lose much of their strength or simply disappear. One alternative consists in placing the infant under the care of an institution established to foster unwanted children. Another, better, alternative obtains when the child is a suitable candidate for adoption. In this case the good the infant will bring to whom might adopt it must be added to the value of the infant's life in our moral calculus. If such options exist at the time the infant is born, it would be impermissible to kill it[227].

We can conceive of cases, though, in which killing infants with a life worth living might be justified. The first kind of case concerns disabled infants. Consider the following situation imagined by McMahan[228]:

> "Suppose, for example, that an infant is discovered to have a defect that could not have been detected prenatally, and suppose further that the defect is sufficiently serious that it is unlikely that anyone can be found who would be willing to adopt the infant. If the interests of the parents that favour infanticide are strong –for example, the burden of caring for the

---

closest to the child do not want it to live. As an infant can be adopted by others in a way that a pre-viable fetus cannot be, such cases will be rare"; see also Singer (1995, p. 211; 2011, p. 166). Warren (1984, p. 117) seems to be of a similar mind.

[227] In a recent article, Giubilini & Minerva (2012) argue that even if adoption is a feasible alternative, a woman has sufficient reasons to kill her newborn infant if giving up the child would damage her, perhaps because the woman "experience[s] serious psychological problems due to the inability to elaborate the loss and to cope with grief". But suppose that Selgelid (2012, p. 53) is right in characterising Giubilini & Minerva's general moral outlook as preference utilitarianism. Then, as he suggests, it is not clear that preference-satisfaction maximisation shall always justify infanticide in the cases Giubilini & Minerva have in mind. Indeed, the woman's preference not to give up her child must be balanced, for example, against the preferences of those who would adopt it, and it is at least conceivable that sometimes the latter shall trump the former. But they ought to say more about their conception of the moral value of preference satisfaction in order for a sound evaluation of their argument to be performed.

[228] McMahan (2002, p. 359), but a similar point is made in Singer (2011, pp. 162-163) when discussing our reasons to kill a haemophiliac infant, under the assumption that merely conscious beings are replaceable.

> defective child would be so great as to preclude the possibility of their having a further, healthier child- it is not unreasonable to believe that they could outweigh the infant's time-relative interest in continuing to live."

In this case, two kinds of reasons count in favour of killing the child. First, the interests of the parents, who shall be deprived of a great good in their lives (having the healthier child) in exchange of having to bear a great cost (caring for the defective child). Second, the impersonal value of the life of the possible healthy child, which shall be lost to the world unless the defective child is killed. If we believe those are sufficient reasons for killing a foetus, we ought to believe that in a case such as this, in which no other options are available, they are sufficient too for killing the infant. As I pointed out some paragraphs ago, if we refuse to accept this conclusion, then we must either reject the view on the moral status of the foetus that implies it, or accept other normative beliefs (about the significance of birth or viability) which are probably false.

There is another kind of case in which infanticide might be justified. Sometimes killing an infant, even a healthy one with a life worth living, might be necessary in order to save the lives of others. McMahan believes that the following would be one such situation[229]. Imagine that there are three five-year-olds who will soon die unless they receive some transplants. If they survive, their lives will be worth living. In the same hospital there is a newborn infant who has the right tissue type for the transplants. All three children can be saved if the transplant is performed. The infant, though, shall die in the process, thereby losing a life which would have also been worth living. Suppose that no one else is worse-off because of the infant's death. Its single mother died in childbirth, and the infant has no known relatives. Also, there are other children available to satisfy other people's desire to adopt. In this case, it would seem that the prudential value of the three children's lives gives us reasons to kill the infant which are stronger than those against killing it provided by the time-relative value of the infant's life. If this is the case, McMahan reflects, we appear to have at least sufficient

[229] McMahan (2002, p. 360).

reasons to kill the infant. In other cases, in which, for example, our reasons in favour of saving other beings' lives are even stronger, our reasons to commit infanticide might be decisive.

It is possible that these cases engage the reader's emotions in such a way as to dispose her against their justifiability. This might be especially true of the last case, in which an infant's death is used as a means to save a few children. It is not difficult to see how there may be nothing wrong with killing an anencephalic infant, or an infant whose life will be worth ending. These things fit quite nicely with a moral concern for the well-being of all sentient creatures. But how can killing infants with a life worth living possibly be in accordance with such moral concern?

First, we ought to remember that on these views a decision to commit infanticide must be reached after taking into consideration the well-being of all those affected by it, including the infant's. Second, it must be stressed that none of these views implies that infanticide is easily justifiable. On the contrary, it is seldom permissible, for it is seldom true than we can expect the death of an infant to be *necessary* for the promotion of the well-being of other people, either present or future. Moreover, even in those cases in which it is permissible, it might still be criticisable. Suppose that some parents are expecting a child and had decided that if it suffered from a particularly serious condition, detectable before birth, it would be best to kill it, even if its life would be worth living. Assume, further that their belief is true, but that they neglect to submit the pregnant woman to the appropriate tests. After birth, it turns out that the child suffers from the dreaded condition, and it is thereby killed. Even if this is a permissible instance of infanticide, it might be true that the reasons against killing the child were weaker before birth. This would be so if we accepted a time-relative value account, for typically foetuses are more weakly related to their future lives in the ways that matter than infants are. It might be true as well if the parents and other participants in the killing are worse-off than they would have been if the child had been killed while still inside the womb –perhaps it

would have been less distressing, or would have elicited less intense grief or guilt[230]. All these things matter for practical reason, for (at least) one of our rational aims is to bring about the best possible outcome.

In the third place, it must be born in mind that our raw intuition against infanticide is methodologically suspect. If we feel upset about the possibility of killing an infant as a means of saving a few lives, we ought to consider whether that feeling would be aroused in similar cases where the victim is either a human foetus or a non-human animal of high psychological capacities (though not high enough to qualify as a person). Perhaps, as McMahan suggests, this unease may stem from our recognition of the existence of a deontological "constraint against harmful using", which would give us decisive reasons, except in extreme circumstances, against *opportunistic killing/letting die*, that is, exploiting someone as a resource in a way which the agent reasonably believes that will eventually lead to that one's death[231]. Suppose that we are ready to accept, as indeed many people seem to do, that there are cases in which we may have sufficient reasons to opportunistically kill a baboon or chimpanzee in order to use their organs to save the lives of a *few* young children. We do not think, then, that a deontological constraint against harmful using applies here. On the views I consider in this section, if there are any differences between our reasons against killing a healthy baboon or chimpanzee and our reasons against killing healthy late foetuses and infants, these are mainly differences in degree, owing to the typically worthier future the latter have. If that is the case, it would be reasonable to claim that killing late foetuses and infants is permissible in fewer cases of this kind than it would be to kill an animal of high psychological capacities. But it would be inconsistent to claim that only in extreme cases we may permissibly harvest organs from healthy foetuses and infants with the aim of saving some children. Why should this deontological restriction apply in favour foetuses and infants but not in favour of

[230] See McMahan (2002, p. 343; 2013, p. 280); Singer (2011, p. 166).

[231] See McMahan (2007, pp. 9-11), whose main argument I follow in the remainder of this paragraph.

chimpanzees? Unless we embrace implausible views on what constitutes a morally relevant attribute –species, viability, birth–, any restriction that may make foeticide and infanticide impermissible in these cases, ought to apply to the killing of an animal with *higher* psychological capacities. Under the fairly plausible assumption, that is, that, provided that such things as deontological restrictions exist, their occurrence is related to the psychological capacities possessed by the beings which shall be affected by our agency.

Hence, that a view favourable to infanticide is implied by some of the theories I have surveyed in this chapter is not an absurdity that warrants its refusal. Though the normative intuition that infanticide is on a par with homicide is widespread, we have serious reasons to doubt that it tracks the moral truth. Because it is methodologically suspect, the fact that we possess this intuition does not give us epistemic reasons for believing that infanticide is as wrong as we pre-theoretically feel inclined to claim. On the other hand, each of the theories at issue benefits from independent support, on which I have elaborated in the several sections of this chapter. If we believe any of those theories, then only at the price of inconsistency it is possible to escape a conclusion favourable to infanticide. Of course, inconsistency is a fault to be avoided. Thus, this deeply ingrained rejection of infanticide is one aspect of our practical attitudes that we ought to revise.

CHAPTER IV
CONCLUSION

§1. *On the aim of this chapter.*

In this chapter, I shall present the overall conclusion of my thesis. I will begin by briefly restating the reasons why the *Future-of-Value Account* is the relatively best view contrary to foeticide. Then, I shall provide a summary of the ways how it can be rejected. Finally I shall remark how even though this account conceives of foeticide as a species of homicide, that is not what we ought to believe.

§2. *On the relatively best view contrary to foeticide*.

The *Future-of-Value Account* states that our reasons not to kill a person are very strong, on the assumption that her future life is worth living. This is so because these reasons are given by facts about the prudential value the victim's life would have if she were not killed, and this value is conceived as time-neutral. Thus, our reasons not to kill a foetus are similarly strong, assuming that it shall become a person with a life worth living (CH. II, **§6**).

There are two main considerations which favour this account over its anthropocentric rivals, as I discussed in CHAPTER II. You may remember that *Anthropocentrism* claims that (CH. II, **§2**):

(1) The mere fact that a living organism is a human animal is always valuable.

(2) It is valuable in that, except in extraordinary circumstances, it gives us decisive reasons not to kill that organism.

In the first place, the *Future-of-Value Account* appeals to what is plausibly a normatively relevant fact, that is, how good or bad a life is for the one who lives it. Contrariwise, all versions of *Anthropocentrism* appeal to an attribute, species membership, for whose rational significance no persuasive argument can be made. Thus, we saw that membership in the human species cannot be defended as either a basic or a derived practical principle about what is impersonally valuable (CH. II, **§3**). I also argued that attempts to ground the personal value of our lives in species membership fail (CH. II, **§4**).

In the second place, anthropocentric accounts imply a series of implausible normative judgements which the *Future-of-Value Account* happily avoids. Impersonal versions of Anthropocentrism, for instance, subordinate personal considerations about our well-being to the impersonal value which they attach to all human life (CH. II, **§3**). In this way, it would have to be concluded, for example, that the life of unconscious human animals is as valuable, in the way that matters, as that of human animals with lives worth living. It would also have to be accepted that the aim of preserving our existence gives us strong reasons even against rational suicide. Those who find these implications highly implausible have weighty reasons to reject this version of Anthropocentrism. Since other versions of this view, which conceive of the value of our lives as personal in nature, also lead to similar conclusions, we also have similarly weighty reasons to reject them (CH. II, **§4**).

The *Future-of-Value Account* is, then, superior to other attempts at defending a view contrary to foeticide. Nevertheless, it is not the most plausible view about this practical problem.

§3. *On the more plausible views about foeticide*.

As I explained in CHAPTER III, there are very strong grounds on which to reject the various claims that constitute the *Future-of-Value Account*. These claims are (CH. II, **§6**):

(a) The fact that dying would be bad for the one who dies gives us reasons not to kill her.

(b) The badness of death for the one who dies at a certain time is determined by the net value her life would have were she not to die at that time.

(c) One can *only* be deprived, in the normatively significant sense, of a future that is *her own*.

To which we ought to add:

(d) We are essentially human animals.

(e) Identity is what matters in survival.

First, it must be clear that facts about personal ontology and numerical identity over time are not what matter in survival (CH.III, **§3**). Numerical identity is, of necessity, a categorical relation which excludes branching. Thus, it cannot account for the intuition that we may be more or less related with our future in the ways that have normative significance; or for the intuition that, for practical purposes, we can survive a situation like *Division*. That does not mean that these metaphysical issues are uninteresting. Ontology teaches us how to determine how many things there are in the world and in what manner they exist. Knowing the truth about quantities and qualities is instrumental in order to learn how to properly describe the practical dilemma of choosing whether to kill

a foetus. Because Lockeans and Animalists disagree about the kind of thing we are essentially (CH.III, **§2**), they disagree about what is the most accurate way of describing this dilemma. On the one hand, if we believe that we are essentially some kind of animal organism, we ought to say that, for the greatest part in its development, killing a foetus consists in killing one of us. It is disputed, though, whether organisms are real existents and, if so, what manner of integration there must be among biological material for it to become one (APPENDIX A). We could believe that a foetus does not become an organism until it loses its capacity for twinning. Then, killing a foetus before the first days of its third week of development consists in, strictly speaking, preventing one of us from existing, but not in killing one of us. Alternatively, we could believe that foetuses become organisms earlier in their development, even right after fertilisation. Then, we ought to accommodate our description accordingly. On the other hand, we might believe, as the Lockeans do, that we are some kind of psychological entity, such as a mind or a person. If that is so, we ought to consider that it would be mistaken to describe foeticide before the eighteenth week of foetal development as the killing of one of us, for it is then that the foetus arguably acquires the capacity for consciousness (APPENDIX B). On some versions of the kind of psychological entity we are, mentation suffices for our existence, so that from that week onwards, foeticide is rightly conceived of as the killing of one of us. On some other versions, some complexity of mental capacities or content is also necessary, so that foeticide may never be described as killing one of us, but always as preventing one of us from existing. I have not attempted to produce criteria for identifying the correct conception of an organism, or even to determine whether such a thing exists. Neither have I attempted to disprove Animalism in the face of the Lockean views, which is the most appealing view to me. Nevertheless, if it is true that identity is not what matters, then whatever the outcome of this discussion is, it is inconsequential for practical reason.

Contrariwise, what is indeed consequential for practical purposes and what does matter in survival is some kind of psychological relation (CH.III, **§3**). Thus,

it is only by being psychologically related to some future state of affairs that its obtainment can be good or bad for a conscious being. Two such possible relations may be distinguished, namely, continuity of psychological capacities and continuity of mental content. We may assume for the sake of the argument that continuity of psychological capacities has normative significance, and even that its obtainment is necessary for a future to be prudentially valuable. This is a fair assumption since it strengthens the views contrary to foeticide. For in conceding this, it is accepted that by the acquisition of psychological capacities around its eighteenth week of development, the life of the future person whom the foetus will become can gain value for it.

Nevertheless, even if such normative relevance is granted to continuity of psychological capacities, cases such as that of *Valjean-Madeleine* engage the intuition that what fundamentally matters in survival is continuity of mental content. In the case that numerical identity was the relevant relation between a sentient being and the stuff that composes its life, it could not vary in degree. At any moment in its existence said being would be related in the way that matters in exactly the same manner with any event in its life. Our judgement of the personal value which any such event would have for it would necessarily be time-neutral. That is, for any life-event, the value judgment of it cannot vary because of a variation of the moment in the being's existence when it is performed. A being, however, can be more or less intensely psychologically related with an event in its life. Thus, the value judgment of such an event may vary with the moment in that being's existence when it is performed, for at different moments a being can be psychologically related with its future more or less intensely. Hence, our judgements of the personal value of a life-event must be time-relative. Persons such as us, i.e., self-conscious entities with a highly integrated mental life over time, are very intensely psychologically related with most events in our existence. Consequently, the time-relative value our future has for us is very similar to its time-neutral value. As a being becomes less intensely psychologically linked to its future, the time-relative value it has for it diverges from the time-neutral one by progressively diminishing.

Now, foetuses, by the time when they acquire the capacity for consciousness, are very weakly psychologically related to the stuff that composes their future, and continue to be so throughout their development in utero. Howsoever high the time-neutral prudential value of their life might be it ought to be acknowledged that its time-relative value is exceedingly low.

Let us assume, first, that our main reasons against killing a sentient being are provided by facts about the prudential value of its life. Let us also assume, for the moment, that practical reasons are object-given and value-based (CH. I, **§2**). It should be concluded, then, that these are absent prior to the eighteenth week of foetal development. Before that time, whatever reasons for or against foeticide we may have shall be provided by the impersonal value of facts about the future life of the sentient being which the foetus would become. Additionally, it should be concluded that after that week there are reasons for and against killing it given by the personal value its life holds for it. Even if such reasons exist, however, they must be relatively weak indeed. This must be so because the relation whose obtainment determines whether the value of the event is personal instead of merely impersonal holds to a very low degree. Thus, because of the minimal intensity of the personal value at issue, the strength of the practical reasons with which it provides us is very similar to that of the reasons we would have if they had been given by an impersonal value. This is the sense in which, under these assumptions, even though our reasons for or against killing a foetus late in its development, as given by the value of its future life, are *stricto sensu* personal in character, they operate *as if* they were impersonal.

Assume now, on the contrary, that practical reasons are subject-given and desire-based (CH. I, **§2**). An event in the life of a sentient being can only be valuable for it, then, if its happening is the object of an appropriate desire of that being (CH.III, **§4**). Therefore, continuing to live shall be prudentially valuable for

a particular being just in case that it possesses a desire to live or whenever continuing to live is necessary for the satisfaction of other desires it possesses.

The central issue is the characterisation of the relevant instrumental desire to live. If it were justified to judge the prudential value of a sentient being's life with a time-neutral standard, it could be argued that from the moment when a foetus first experiences desire-like states, its whole existence becomes instrumentally desirable for it, as a necessary condition for the continued enjoyment of those events it values. Nevertheless, as it has been discussed, the prudential value of a life merits a time-relative, and not a time-neutral judgement. It is possible, however, to produce desire-based accounts of the badness of death which are compatible with the claim that what matters is some kind of gradual, psychological relation and that the value of the stuff that composes a being's existence is therefore time-relative. These are accounts which temporally restrict the extension of a creature's life which may have instrumental value for it to as far in the past and to as far in the future as said creature's mental states –such as memories or expectations– allow. When such a creature is killed, it is the prudential value of that temporal slice of its life what provides us with our main reasons against killing it. On the contrary, the value of the future existence it would have lived which is not covered by the mental states possessed at the time when its death is considered must be impersonal. There are several theoretical possibilities regarding the identification of the time-slice which ought to be the subject of prudential valuation. One option consists in merely requiring the presence of mental states which somehow involve the notion of being temporally extended, such as having expectations for the near future. Another option, on the other side of the spectrum, consists in additionally requiring the belief that one is a temporally extended entity. Different sets of cognitive capacities allow for different lengths in the temporal extension, as here understood, of a creature's mental life. A person's mental life typically extends very far into the past and is able to harbour expectations and develop plans for her furthest future. Thus, our reasons against killing persons, as given by the prudential value of their lives, can be very strong. Yet a sentient foetus' mental

life extends but slightly, if at all, into the future. Under our current assumptions, this entails that the value of such future life must be impersonal. But that is also true *a fortiori* of the pre-sentient foetus. Hence, it must be concluded that at any point in its development, our reasons for or against killing a foetus, as given by the value of its life, are *stricto sensu* impersonal in character.

Prudential value can be conceived of as something to be promoted, thus bringing about a requirement of outcome-value maximisation. As I made explicit, it has been assumed that our main reasons against killing are provided by this kind of value. This is part of a more general outlook which considers that our most important practical reasons are those given by facts about personal value. By concluding, in the different ways referred to in the previous paragraphs, that our reasons for or against foeticide are either *stricto sensu* or *as if* impersonal, we accept that the value of foetal life plays a subordinate role in our practical deliberation, which is dominated by the well-being of the other parties affected by the decision.

There is, however, an alternative way of arguing for that subordinate role of the value of foetal life. This consists in justifying the absence of a deontological constraint against killing foetuses which is present against killing other sentient beings with more complex psychological capacities. One manner of doing this consists in conceiving the prudential value of a life as something to be respected, thereby bringing about some sort of constraint against killing which is more or less resilient to demands of outcome-value maximisation (CH. III, **§4**). But according to the plausible views about the value of foetal life that we have considered, it must either be impersonal in nature or very much as if it were so, regarding its intensity. Thus, it must be protected by either no constraint against killing or by an extremely weak one. Alternatively, we could assume that the main reasons against killing one of us are not given by the prudential value of our lives, howsoever conceived. We may suppose instead that these reasons are provided by a requirement to respect an autonomous choice not to die (CH. III, **§5**). Such a requirement would be, as well, resistant against demands of

value maximisation. Yet the foetus does not possess autonomy, so that there is no choice against dying that must be respected. The value of its life must be, then, integrated into the calculus of value that is to be promoted. Again, and for reasons already discussed, such value ought to be considered either impersonal *stricto sensu* or *as if* it was so.

§4. *On what we ought to believe about foeticide.*

None of the views that I have considered claims that there cannot be decisive reasons against foeticide. What these theories do claim, however, is that we would be mistaken in treating these cases as a species of homicide. What matters most in homicide is that a person loses a great good, or that her choices are disregarded, or both. None of these things matter much in foeticide. What is important, instead, is how the existence of that being which the foetus would become would affect the well-being of present or future actual people, as well as how its own value shall contribute to the aggregate of goodness in the world. Since these are also the things that matter most in deciding how the world ought to be populated, the right way of conceiving of foeticide would be as a problem of population ethics. Because, in fact, people are the result of reproductive choices, foeticide is also a problem of reproductive ethics.

Suppose that these views are right in conceiving of foetal life as impersonal, or nearly so. Could it be true that, given the relevant facts, on most occasions people have decisive reasons to let live whatever number of children they happen to conceive? Could it be true that, even on these views, foeticide is seldom permissible? In order to provide a reasoned reply to these questions we need first to determine what is the impersonally best way for the world to be populated. Thus, we need to identify (a) how many sentient beings and (b) what distribution of well-being among them it would be optimific, in this sense, to exist. In the second place, it would have to be established whether what we have most reason to do always coincides with what is impersonally best.

Perhaps regarding certain decisions it agrees with practical reason that an agent chooses what is best for those she is partial to (including herself) over what is impersonally best. Then, we ought to find out whether any reproductive choices belong to this kind of decisions. Alas, although these are the questions at which my reflections point, it lies outside the scope of this thesis to discuss them.

So, what do we have most reason to believe regarding the moral problem of foeticide? On different, incompatible assumptions it is possible for moral philosophers to agree that foetal life is impersonally valuable, or at least that it is very much as if it was so. For those of us who seek an answer to what to believe and what to do about foeticide, this is encouraging. It shows that, in spite of the underlying theoretical differences regarding the nature of practical reasons or the truth about the wrongness of killing persons, what appear to me as the most plausible views can reach a broad consensus on the strength of our reasons for and against foeticide, and about the most illuminating way of understanding this practical problem.

APPENDIX A
ON WHETHER AND WHEN A FOETUS BECOMES A HUMAN ANIMAL

When do human animals begin to exist? This is the question that has engendered one of the most byzantine discussions in this literature. Philosophers take the soundness of their arguments to hinge on obscure minutiae about embryonic development and, consequently, fill pages with their interpretations of biological facts. I would be loath to introduce the reader to this issue save for its relevance to the morality of so-called chemical abortion, that is, foeticide by hormone ingestion. Though these procedures are characterised as a form of 'emergency contraception' by medical authorities, as opposed to being abortifacients, it would be wrong to believe that there is a substantial discrepancy here. What happens is that these same medical authorities define pregnancy not as beginning with fertilisation, but about five days later, when the fertilised egg implants itself into the uterine wall. Whatever prevents implantation, then, is considered a contraceptive, rather than an abortifacient. But when the implantation of a fertilised egg fails, it dies. Since these hormonal 'contraceptives' may produce sometimes their effects precisely in this way, it is worth asking whether what has been killed is a human animal or not[232]. If the answer we ought to rationally give is negative, we will have found that views contrary to foeticide will prove unable to object to chemical abortion, in general, and the use of emergency contraceptive pills, in particular.

Since nothing is ever easy, whereas some writers are convinced that human animals begin to exist right after fertilisation, some others have serious doubts about it. Yet before reconstructing the different positions regarding the moment when human animals begin to exist, let me briefly acquaint you with the facts

[232] See Trussell & Raymond (2012, p. 8) for the medical definition of pregnancy and the mechanisms through which emergency contraceptives produce their effects.

about early human development, from fertilisation until the beginning of the third week afterwards, when certain events –generally accepted as relevant for our purposes– take place[233]. In the following paragraphs I will take the liberty of referring to the prenatal human entity during this early period as the 'embryo', introducing the proper stage-specific denominations when necessary.

Fertilisation begins when the sperm penetrates the egg. Some thirty hours later, out of their combination emerges a two-celled entity called *zygote*[234]. Let us examine more closely what a human zygote is like. As I just mentioned, it begins its existence as two adhering cells. These cells –the *blastomeres*- are surrounded by a protein shell called the *zona pellucida*, which they have inherited from the egg. Until the eight-celled stage, blastomeres form a loosely arranged collection inside the zona. Then, a process of compaction begins as the blastomeres continue to divide, with the inner cells of the zygote becoming much closer together than its outer cells. Around the fifth day after fertilisation the zona pellucida has disappeared, and the embryo, now referred to as a *blastocyst*, has experienced some remarkable morphological changes: the inner cells have merged together into a mass, attached to the inside of a cavity whose walls are formed by the outer cells. Also, we can observe the very first hints of functional differentiation, for whereas the outer cells (which constitute the majority) will develop into the placenta and other supporting structures, it is only the inner cell mass that will give rise to the embryo proper. In fact, it is from this inner cell mass, and not from the outer wall, that embryonic stem cells are harvested for therapeutic purposes, due to their pluripotency, that is, their capacity to develop into almost any cell type.

[233] I base my embryological claims on Sadler's account (Sadler, 2012). I have drawn from pp.33-39 for an explanation of fertilisation and the pre-implantation development of the embryo; from pp.43-47 for the process of implantation; from pp.51-52 for the formation of the primitive streak and from pp.110-113 for some facts about twinning.

[234] The phrase 'single-celled zygote' is quite popular in the literature, but there is actually no such thing. The male and female DNA synthesize and immediately afterwards the fertilized egg divides normally, becoming a two-celled zygote. At no time there is a *single* cell with a complete set of human DNA in its nucleus directing its metabolism (Sadler, 2012, pp. 34-35).

The blastocyst then begins to embed itself in the woman's womb. This is the process called implantation, and it will last for about a week, until around the twelfth day after fertilisation. During this time, the functional specialization of cells both in the outer and inner masses gains pace, with the latter acquiring the shape of a flat disc composed of two superimposed layers. It is in the layer of greatest interest to us –the *epiblast*- where the momentous event of *gastrulation* takes place. By the beginning of the third week of development a narrow groove appears on the epiblast's surface. This groove (the *primitive streak*) consists in a group of cells specialized in synthesising a protein that controls the migration and specification of all the other cells in the epiblast, which is transformed into a three-layered structure. It is from this structure that all the tissues and organs in the foetus will arise.

Now that I have familiarised you with the technicalities, I can describe the competing opinions about the beginnings of human animals more precisely. On the one hand there are those who believe that humans begin to exist very early in pregnancy[235]. These writers claim that human animals begin to exist right after fertilisation as two-celled zygotes. At first sight, this may seem plausible. According to some authors, though, this impression disappears once we look more closely into the pertinent biological facts[236]. These writers deny that we

[235] Such as Carter (1982, pp. 93-95), Finnis (1973; 2011b), Kaczor (2011), Lee (2010), Newton (1975, pp. 334-335), Perry (2007, p. 58), and, of course, the Catholic Church: "procured abortion is the deliberate and direct killing, by whatever means it is carried out, of a human being in the initial phase of his or her existence, extending from conception to birth" (*Evangelium Vitae*, 1995, par.58).

[236] Among these there are those who deny that we have decisive reasons to believe that zygotes are human animals. These are such authors as Becker (1975), DeGrazia (2005, pp. 245-254), Marquis (2007b, pp. 196-205), McMahan (2002, pp. 24-29), Meyers (2010, p. 71), Olson (1997b, pp. 89-93), Steinbock (2006; 2007; 2011), Singer (2011, pp. 144-146), Wiggins (2001, p. 239) and, also, Thomson (1971, p. 48), though she gives no reasons for her belief due to her argumentative strategy. Others, though, like Devine (1984) or Warren (1997, pp. 202-205), have serious doubts about the zygote's membership into our species or, like Glover (1977, p. 121), find it unclear how to settle this dispute.

have decisive reasons to accept that a new human animal has begun to exist right after fertilisation, and yet they disagree as to what exactly we ought to believe in this respect. At the end of the day, I believe that the differences among these different positions can be explained as a difference in their conception of an organism. They disagree about the *degree of functional integration* that must obtain among a number of cells for it to be true that they have a *common* metabolism.

An examination of one of these views will shed light on the relevance of a shared metabolism. As I said, some favour what I will call

> *Weakest Integration*: cooperation of the sort that exists right after fertilisation among embryonic cells is a *sufficient* condition for enough integration among them.

This position, though not entailed by Anthropocentrism, is common among its supporters[237]. Why consider fertilisation as the best candidate for a cut-off point between non-existence and existence? It certainly seems most conspicuous in one respect –it appears to be a point of radical discontinuity just before the start of a developmental process that will be both continuous and gradual. Thus, we find two distinct entities –the egg and the sperm– which, fusing, cease to exist and bring about the existence of a new two-celled entity. From then on, it is said, we can trace the development of that entity step by step, that is, cellular division by cellular division, until birth. Moreover, there are no interruptions in this process of ongoing division and functional specialization of cells and it is not as if great morphological or functional changes occur all of a sudden. In this respect no stage in this process is more dissimilar to its adjacent stages as the creation of the new entity is to the fusion of the two others that created it.

[237] In reconstructing the case for the end of fertilisation as the beginning of a new human organism I am chiefly following Finnis (1973, pp. 144-145; 2011a, p. 281), Lee (2010, pp. 71-76) and Kaczor (2011, pp. 102-105). Other authors that share this view usually appeal to the same facts as these.

If this was all that there is to be said for fertilisation, it would be a poor argument indeed. The fact that a process is gradual or each of its constituting stages inconspicuous does not imply that any distinction we make among them is arbitrary –that will only be so if *no* change during the process is relevant[238]. Imagine that any temporal stage in a process of development is contiguous to another such stage. Imagine moreover that these temporal slices can be cut so thinly so as to exhibit little difference with what goes on in the slices immediately prior or subsequent to them. None of these facts precludes the possibility of identifying some set of contiguous temporal slices (some period of time) such as that the questions about what exists *before* and *after* it merit different answers. Even if it is not clear what we ought to answer if asked about what it is that exists *during* that period. Thus, the fact that gestation is a process of gradual development does not necessarily preclude the possibility that human animals begin to exist at some time during it. It is not enough, then, for supporters of this position to remark on the *relative* saliency of fertilisation –they have to show that *what it consists in* justifies the claim that a newly created zygote is a human animal.

What appears to be important about fertilisation is that a new combination of DNA is produced. This is why. Let us assume that an organism is a material object with a *life* of its own. An organism's life is an event which consists in a sum of complex chemical processes. These processes, collectively called *metabolism*, are directed to maintaining the structure of the organism, which otherwise would decay. The multiple parts of the organism are arranged and cooperate in such a way so as to constantly replace the particles of matter that

[238] This point is made by Boonin (2003, pp. 38-39). He also notes that, since fertilisation is itself a complex process of gradual gamete dissolution and cellular fission we ought to expect similar problems in determining when exactly it ends. This same difficulty is highlighted by Tooley (2009, pp. 44-46).

constitute it, repair them and adjust its functioning to the environment[239]. And the basic information regarding how the chemical processes of an organism are to operate is made of DNA. It will be present in each of the organism's cellular nuclei. This DNA will determine, as it were, *from within*, the growth and development of the organism. Well, the zygote –we all agree- will develop into a human animal in a few months' time. Its DNA is present since its very beginnings right after fertilisation and it is the designing force behind most of the changes that will ensue. Thus, the zygote is already envisaged as a "self-developing and self-integrated whole", as opposed to a "mere collection of human cells"[240]; and possessor of the "internal resources and active disposition' to fully mature, just in need of a 'suitable environment"[241]. Now, we had assumed that this is what being an organism is all about. If that new organism is not *a* human animal and *the same* human animal as that which will exist in the future, then, what is it? The soundest opinion, it is claimed, is that we all began our existence as two-celled zygotes.

How has this position been challenged? First, it can be argued that the fact that several organisms engage in a cooperative process the result of which is a new life-form doesn't necessarily show that during said process the cooperating organisms already constitute that life. A fine example of this is cellular fission[242]. Cellular chromosomes begin to divide and they cooperate with other subcellular organisms to rearrange the matter that composed the cell until it physically splits in two. Before this process starts we have one organism, the ancestor cell. Once it is finished, two new organisms are present. What about the intervening period? One possible answer is that during this period there is no *single* life present, but that of the various subcellular organisms that for the time being are part of no larger structure. We may be reluctant to go as far as to

[239] This is, succinctly stated, the metaphysical conception of organisms defended by DeGrazia (2005, p. 245), Olson (1997b, pp. 126-140; 2007, pp. 27-29), but first and foremost by van Inwagen (1990, pp. 142-158).

[240] Kaczor (2011, p. 105).

[241] Lee (2010, p. 73).

[242] I borrow this example from van Inwagen (1990, pp. 150-151).

claim that and rather say that until its physical division the ancestor cell continues to exist. It does not seem plausible, however, to claim that the daughter cells exist from the beginning of their ancestor's fission process. That would require the existence of two separate metabolisms during a time it is not obvious there is even one.

So only because there is some cooperation among the zygote's cells we should not assume that they compose an organism. Therefore, some believe there are reasons for

> *Weak Integration*: the presence of functional differentiation among embryonic cells is a *sufficient* condition for enough integration among them.

You may remember that the first differentiation among the embryo's cells occurs by the fifth or sixth day after fertilisation. By then we can clearly distinguish between its outer cells and its inner cell mass and we know that their future development will be radically dissimilar –the inner cells will develop for the most part in the embryo proper, whereas out of the other cells supporting structures will emerge. What *Weak Integration* claims is that before that differentiation has taken place the embryo cannot be considered an organism, much less a *human* organism. Indeed, it has been suggested that the very early embryo is merely a "virtual object", that is, not a physical object but only existing (in a loose sense) because of how we ordinarily speak[243]. We have assumed that a group of cells

[243] The earliest presentation, that I know of, of this ontological argument can be found in van Inwagen (1990, pp. 152-154), but has also been put forward by Olson (1997b, pp. 92-93). I borrow the phrase "virtual object" from van Inwagen (1990, p. 112). Of course, what counts as a virtual object will depend on which categories of 'real objects' we are willing to admit. Van Inwagen's ontological inventory, though, includes only physical simples and organisms. If artefacts (such as chairs or societies) count as real objects, then the zygote might really exist after all. But even if organisms themselves are considered artefacts, the point is still that because of the conception of organism *Weak Integration* defends, zygotes cannot be considered one.

compose an organism –as opposed of being merely spatiotemporally contiguous– only when they are all trapped in the same metabolic process. Arguably that is not the case here. Picture the zygote –a bunch of undifferentiated cells in a protein shell. The fact that we happen to have a collective name for all these cells ought not to beguile us. For one, it is not the case that we usually can rearrange the parts of an organism without affecting the outcome of its processes, or without it trying to restore its original structure, but that is just what happens with the embryo's cells during this period[244]. Also, it is not as if by denying that he zygote is an organism we are stuck with an ontological mystery –a perfectly sound alternative explanation of what it is can be found. Consider again what happens when any cell in our body divides itself. By fission it ceases to exist and is replaced by two new numerically distinct cells. It doesn't even occur to us to suggest that the mereological sum of the succeeding cells compose a single entity, a new organism. Or that the organism that was the ancestor cell did not cease to exist but persists as a two-celled organism. These possibilities are precluded by the fact that the descendant cells share no common metabolic process of their own but rather are integrated into our metabolism. This teaches us something about what happens to a zygote cell each time it divides –it ceases to exist and two new entities come to occupy its place. According to this view, then, the zygote is not a real object that persists through time, gaining new parts as its cells divide. During this period 'zygote' is just the name we give to a number of transient single-celled organisms inside of a membrane.

[244] Cf. van Inwagen (1990, p. 157), regarding a later stage of embryo development but applicable to the zygote too: "It no more has metabolism than does a human pyramid at the circus: The individuals that virtually compose it metabolize individually, but it does not. Nothing directs is growth: Its growth is the sum of the uncoordinated growth and fission of its component cells. (Uncoordinated in the sense that none of the component cells adjusts the rate of its growth or the time of its fission in response to environmental conditions caused by the growth or fission of the other component cells –but no doubt the activities of the component cells are coordinated by a "pre-established harmony")".

In addition, some facts that we know about *monozygotic twinning* give us reasons in support of *Weak Integration*. I suppose we are all familiar with monozygotic twinning (hereby simply 'twinning'). It occurs when the embryo divides, and though it can happen as early as the two-celled phase and as later as the third week after fertilisation, it usually takes place during the early blastocyst stage[245], when the kind of functional differentiation we are considering here is taking place. The result of this is what is commonly called a pair of 'identical twins'[246]. Now, the twinning of a blastocyst is very different from that of a two-celled zygote. In the latter case each cell goes its separate way and develops independently. In the former case that is not so, for it is not the whole blastocyst that splits. The outer cells remains whole, only the inner cell mass divides. That means that each twin can trace its individual existence only as far back as the time of inner cell mass division. When coupled with the already mentioned fact that the outer cells of the blastocyst develop into the placenta and other supporting structures, we are left with the strong suggestion that, in fact, the individual story of each human animal can only be traced as far back as the emergence of the inner cell mass.

Nevertheless, some other facts about twinning may give us qualms regarding *Weak Integration*. I just suggested that each twin can trace its existence only as far back as inner cell mass division. Why can it not be true of a *pair* of twins that they began to exist just after the outer/inner cell mass differentiation?[247] Imagine that human animals John and James are monozygotic twins. If we claimed that they *both* began to exist at the time of such differentiation we would be claiming that in the past they were the *same* object –the same mass

[245] I have based my account of twinning and its varieties, again, on Sadler (2012, pp. 110-112).

[246] Fraternal or dizygotic twinning is a very different phenomenon from that of monozygotic twinning. It occurs when two female eggs are simultaneously shed during the same menstrual cycle and each is fertilized by a different sperm (Sadler, 2012, p. 110). The resulting children need resemble each other no more than non-twin siblings do.

[247] I will be just reproducing here a common argument in this literature. See, for example, DeGrazia (2005, pp. 246-248), Kaczor (2011, p. 103), Lee (2010, p. 75), McMahan (2002, p. 26), Singer (2011, p. 145).

of inner cells. Thus, we would be claiming that they were *numerically* identical to it. We must bear in mind, though, that numerical identity is transitive. If we claim that an *x* and a *y* are one and the same object as *z*, then it necessarily follows that *x* and *y* are the same object. Thus, in this case we would imply that John and James are numerically identical to each other. But they clearly are numerically distinct organisms –sometimes John is ill when James is perfectly healthy, one is asleep while the other is awake and, more tellingly, one of them may continue to live when the other has died. Therefore it must be false that they *both* were numerically identical to an earlier inner cell mass right after an embryo experienced outer/inner cell mass differentiation. Yet could it be that *one* twin, but not the other, is numerically identical to such mass? There is no reason to suppose that, since both are qualitatively identical in every important regard and equally continuous with it. A further alternative is even more contrary to our understanding of what it is to be an organism of our species –to suppose that James and John are proper parts of a *spatially scattered* human animal. The most sensible thing to believe, then, is that each began to exist just after twinning occurred[248].

Also, daughter inner cell masses can split again, and again, producing triplets, quadruplets and so forth. It is true of each member of an *n*-plet that he or she can trace his existence only as far back as the time of the division of the ancestor inner mass. If a supporter of Anthropocentrism adheres to *Weak Integration*, she has to conclude that each inner cell mass counts as a human animal, an extremely valuable entity. The problem, though, is that in dividing it ceases to exist. These would be ephemeral entities of our kind that are sacrificed so that two or more others can begin their existence. Now, this is

[248] Meyers (2010, pp. 83-100) tries to apply a similar reasoning to the origin of *chimeras*. A chimera occurs when zygotes or blastocysts produced by dizygotic twinning combine, creating a *single* new entity with *two* different and complete sets of DNA. If we claimed that a chimera is identical with, say, its two precursor blastocysts we would imply that the latter were actually one and the same, which is absurd.

embarrassing for a view contrary to foeticide such as Anthropocentrism[249]. One of its core tenets is that a living human animal is so valuable as to usually give us decisive reasons not to kill it. For that to be so, it must be valuable indeed, and our attitude of regret towards the loss of such great value must be accordingly attuned. We ought to grieve for it, just as we ought to grieve for the death of a newborn baby or a late foetus caused by some unavertable accident.

There might be some, however, who are ready to accept these implications regarding the attitude we ought to have towards cell masses divided in twinning[250]. Yet adherents to Anthropocentrism who are impressed by this *reductio* may want to admit that, at least in cases of twinning, we had better not conceive the ancestor inner cell mass as necessarily members of our species[251]. Thus, they could claim that early functional differentiation in the blastocyst is *necessary* for enough integration, but not sufficient. Also, it could be added, once twinning is no longer possible the force of this *reductio* decays. Thus, it could be claimed

[249] Here I draw upon McMahan (2002, p. 26) and DeGrazia (2005, p. 253). A religious version of this objection ('God could not allow such wastage of embryos') is considered by Lee (2010, pp. 106-107).

[250] Like Oderberg (1997, pp. 270-271), who claims that maybe we ought to grieve the loss of zygotes as we grieve the death of one of us. John Finnis, for his part, accepts that some human individuals die because of twinning and chimerism. In one article he claims that "the division of an embryo into twins or triplets is simply a change from one individual into two; whether the original individual was predetermined to become two, we find at all stages of this remarkable biological process nothing other than an individual or two or more individual human beings" (2011d, pp. 223-224). In another article he even claims that twinning is 'an unusual form of being generated' and that the relationship between ancestor and descendant zygotes is "an unusual form of parentage" (2011b, p. 290). In the same passage, and consistent with this line of reasoning, he conceives of chimerism as "presumably [...] an unusual way of dying". He is aware of the fact that this implies that a "high proportion" of human individuals "never get beyond the earliest stages of existence" but he rejoins that this "is not intolerable to reason" (2011b, p. 290).

[251] Of course, these implications can also be averted by relinquishing the claim that being a human animal is so valuable, but it would be tantamount to abandoning Anthropocentrism.

Moderate Integration:

(a) The presence of functional differentiation among embryonic cells is a *necessary* condition for enough integration among them.

(b) The preclusion of twinning is a *sufficient* condition for enough integration among embryonic cells.

Yet this is an unsatisfactory solution. These conditions for necessary and sufficient integration have been prompted by a desire to avoid unsavoury normative implications, yet these are frail grounds on which to found an ontological (or conceptual) distinction. Such a distinction must be based on facts about how objects (or our thoughts) are, not facts about their value. Though this couple of claims allows us to escape the *reductio*, it will only be at the price of endorsing an arbitrary position, unless we can find some better argument for it.

Well, *what* could that be? The argument could go like this. Paradigm members of our species do not divide (unlike, say, starfishes or some species of worms). The fact that an embryo does shows that it is not one of us, but rather a precursor. As it is usually put, it has not yet *individuated*, that is, it is not yet one object distinct from other objects, whether of its same kind or of different kinds. However, if an embryo does not split but develops into a foetus and then a newborn, and so forth, we can claim that it was a human animal from the moment when the inner cell mass of the blastocyst differentiated from the outer cell mass. Of course, we cannot know in advance whether the embryo will split or not, but the fact that it does is evidence –so could the argument go– that individuation has not occurred and that until that time it never was an individual organism.

The cogency of this argument will at least partly depend on whether twinning is triggered by the environment or by some property of the embryo itself. In the latter case, the argument may prove persuasive. Some embryos have an attribute which others do not. This manifests itself in that some twin when others

develop normally, which fact allows us to say that the latter are individual organisms while the former are not. If twinning was triggered by the environment, that may mean all embryos are equal in this sense. Unfortunately for the proposed argument, the latter seems to be the case. An embryo can be artificially split in the laboratory[252], each of its successors successfully developing into foetuses. This seems to show that *all* of them possess the capacity for twinning. If that is evidence that the embryo is not an individual organism, then until such capacity is lost we have reason to believe that it is not a human animal.

Thus, capacity for twinning may be seen as evidence of non-individuation, of insufficient functional integration among embryonic cells. Can we find further grounds that justify the claim that the blastocyst is not an organism? *Weak* and *Moderate Integration* based their positions on the presence of early functional differentiation among embryonic cells. Yet we know that all our body tissues developed from the inner cell mass, not the outer cells. Should not we look only for the time when functional differentiation appears in the *inner* cell mass, then? Actually, there is a case for the relevance of the outer cells if we conceive the placenta and other supporting structures that will develop from them as organs of a human animal[253]. Unlike most organs, they are temporary. Also they are distinctive in that they begin to differentiate themselves from the rest of the embryo's body much earlier than the other organs. Though odd, it is not absurd, for the need they cover is just temporary and arises before the needs other organs attend to.

Yet even if the placenta and other structures *will be* organs of the embryo, that doesn't mean they *are* so at this stage. For remember that mere spatiotemporal contiguity among cells is not enough for an organism to exist. Neither the fact that the outer cells are somehow coordinated to produce those structures is conclusive, for that was not enough in *Weak Integration* to consider the zygote

[252] Singer (2011, p. 146).

[253] Both Finnis (2011b, p. 290; 2011d, p. 223) and Lee (2010, p. 99) endorse this opinion.

an organism and it cannot be so here where the degree of integration required by the *Moderate* position is being criticised as too low. It ought to be shown that the inner cells adjust themselves to the developments in the outer cells, and vice versa, that is, that both groups share a metabolism. Obviously, they will. But in fact it is not until the *end* of the third week after fertilisation that the outer cells form blood vessels and that these connect with those that, by that time, the embryo proper has developed. And yet, these vessels do not start supplying nutrients and oxygen until the fourth week, when the heart begins to beat[254]. By the time we may consider the supporting structures organs of the embryo, it has been a few days since it lost its capacity for twinning. That happens right after gastrulation, which you may remember occurs in the first days of the third week after fertilisation.

In light of these considerations some authors favour the following position,

> *Strong Integration:* the preclusion of twinning is a *necessary* condition for enough integration among embryonic cells.

Again, we face the question of what would count as *sufficient* integration. Adherents to all the other positions I have reviewed (*Weakest*, *Weak* and *Moderate*) would agree that loss of the capacity for twinning is a sufficient condition for enough integration, since these views consider lower degrees of integration as already sufficient. Some prominent authors who, given my definitions, would count as proponents of *Strong Integration* would also deem the preclusion of twinning as sufficient[255], though opinions are otherwise diverse among writers of this persuasion[256].

---

[254] See Sadler (2012, p. 60).

[255] Like Olson (1997b, p. 89) or Wiggins (Wiggins, 2001, p. 239).

[256] Van Inwagen's opinion is not so clear (1990, p. 154). At one point he claims that "[he] should prefer to think that if an embryo is still capable of twinning then it is a mere virtual object". But in the subsequent paragraph he admits that he simply doesn't know. As lower limit he establishes "the inception of cell differentiation", but without further specification this is problematic, as we have seen. As an upper limit he establishes "the development of a functioning central nervous

After all is said and done, I find it difficult to identify one of the four described views as decisively better than the rest. Also, I cannot say that the degree of enough integration presupposed by any of them strikes me as absolutely implausible. That might be because I am not acquainted well enough with the relevant biological facts. It might be the case, though, that we simply do not know in enough detail what is going on in the relevant stages to give a definitive answer to the question about the beginnings of human animals, or organisms more generally. Yet, of course, all this presupposes that there *is* an answer – some fact of the matter we have yet to discover. Under certain ontological assumptions that is true, e.g., if we assume organisms are real objects. But what if we do not?

Assume that organisms exist just because of facts about how we speak, that they are virtual objects[257]. An organism, then, would not be a thing composed of several other things, its parts –only said several other things would exist, and they would be proper parts of no greater object. That is what was meant when I discussed the opinion that the zygote does not really exist. Blastomeres exist, the particles that make up the zona pellucida exist, they are spatiotemporally arranged in a certain way, and that is all there is to it. Certainly the fact that zygotes exist only nominally does not prevent us from having a moderately

---

system", that is, "about twenty-four days after conception". He "should like to think", though, that the answer could be narrowed down to a moderately small interval. Devine (1984, pp. 35, 36 and 42), for his part, seems to consider the loss of the capacity for twinning as a serious alternative to fertilization for the beginning of human animals (he does not use that phrase, though, but speaks of "human persons" or "human organisms" indistinctively).

[257] See McMahan (2002, p. 29), who holds this view and compares the existence of a human organism with that of a nation. Steinbock (2007, p. 422) explicitly concurs with McMahan, even though in a previous work (2006, p. 27) and in her latest treatment of the issue (2011, pp. 209-210) she simply states that the capacity for twinning makes it, respectively, "dubious" or "not obviously true" that the zygote and blastocyst are human animals. Early in the contemporary abortion debate Daniel Callahan (1970, p. 395) made a similar point after reviewing several positions about the beginning of "human life" (as he called it): "They each represent *a way* of interpreting the data, and not an unreasonable way" (italics his).

sharp answer to the question about its existence: it begins to exist right after fertilisation and ceases to exist when it loses the zona and outer cell/inner cell mass differentiation begins. We have pretty definite criteria about what counts as a zygote. Since its existence is only nominal, though, whenever we assert that 'it has begun to exist' or that 'it has ceased to exist' we are not making claims about what material objects populate the world but only about the proper use of the noun 'zygote'.

The same would be true of existential claims about organisms, with the sad circumstance that, as we have seen, we do not share exactly the same criteria about how to use that term because we disagree about what counts as enough integration. Questions about the existence of an organism would be of the sort Derek Parfit calls empty[258] –different answers to them do not represent different views about *which objects* populate the world, but different *descriptions* of such objects. That need not prevent us from agreeing on many cases about what can be considered an organism (just peep at a biology book) and what cannot be considered so (peep at a geology book). That is, there are instances in which there might be a right description of events functioning as the proper answer to the question, given the concepts we happen to share. But regarding the existence of organisms our disagreement in the relevant criteria would prevent us from agreeing on fringe cases, such as, unfortunately, the beginnings of human animals. In these occasions the empty question would have no true or false answer. Any description would be just as good[259].

[258] Parfit (1984, pp. 213-214).

[259] Becker (1975, pp. 341-345) would claim that this whole discussion misses the point, since he believes that being a human animal, such as we are, requires possession of basic human form and complete organ differentiation. Thus, ascertaining that the multicellular embryo is an organism would not be enough. Yet his main argument for this view is the claim that human ontogenesis is a kind of metamorphosis. But consider the metamorphosis from caterpillar to butterfly. Either they are the same animal or not. If they are, and ontogenesis is like metamorphosis, Becker's argument collapses, for that means that if the embryo is a human organism it must be a human animal. If the caterpillar and the butterfly are not the same animal, his argument may work, provided human ontogenesis is sufficiently alike metamorphosis. But it

I have attempted to give a summary of the current discussion about the beginnings of human animals. One of the most important things that ought to be remarked is that we must shed away the notion that this is entirely a question for biologists. Biological facts cannot determine the answer all by themselves because the extent to which they are relevant depends on our ontological assumptions. If we believe organisms are real objects, then biological facts are most relevant, but we still need a criteria of enough functional integration for their parts, that being precisely the subject of controversy. What normative implications does the claim that organisms do not really exist have? I am sure that of all the possible accounts of the wrongness of killing which entail a view contrary to foeticide some are compatible with the claim that organisms are not really existing entities. Thus, it is possible, in principle, to hold both views. I do not think, though, that the *Future-of-Value Account* in particular can give the reasons against foeticide Marquis intends without the claim that organisms are real existents. As I mentioned, two of Marquis's assumptions are that diachronic identity is the relation with a future that has normative significance and that we are essentially some kind of living organism. If organisms do not exist, then they cannot be what we are essentially, nor furnish us with our criteria of persistence over time. Do organisms really exist or do they not? I do not have the means to settle this question. I am content, though, with granting for the sake of argument that they do, since that is so crucial for Marquis's position.

On the other hand, Marquis's view (as well as the other views I survey in CHAPTER II) is compatible with any of the conceptions of what counts as enough integration among some biological material for there to be an organism. Marquis

---

seems that the latter involves "virtually the complete dissolution of all the organ systems" of the caterpillar [Denton (1986, p. 220), as reported by van Inwagen (1990, p. 155)] whereas none of this takes place during gestation –once it is plausible to say the embryo is an organism (according to whichever criterion you select) it never seems to lose that status. If human ontogenesis is so unlike metamorphosis, I see no reason to have a concept of human animal as demanding as Becker's.

does favour *Strong Integration*[260], and thus he does not believe that one of us exists before twinning is precluded. Yet his argument for this position is completely independent from his account of the wrongness of killing. Hence, it is a mere contingency that many of those contrary to foeticide extend their opposition to the first days of foetal development. One lesson to be learned from this discussion is the following –even if we hold a view contrary to foeticide we still have at least sufficient epistemic reasons to adopt a stance either for or against its moral permissibility during the early stages of pregnancy.

[260] Marquis (2007b, pp. 196-205).

APPENDIX B
ON WHETHER AND WHEN A FOETUS BECOMES CONSCIOUS

In CHAPTER III, I survey a series of ontological and ethical theories which place great stock in our capacity for consciousness. Thus, according to the Lockean views about our transtemporal persistence, the moment when that capacity is obtained marks the earliest possible time when we might have begun to exist. All Lockeans agree that we could not have existed prior to that point. Also, according to several theories about the wrongness of killing, this is the earliest time at which a desire to continue to live can be attributed to the foetus, or the earliest time when it is possible to claim that its life has any time-relative value for it. Throughout my explanation of these theories, I claimed that foetuses gain the capacity for consciousness around the eighteenth week after conception (or the twentieth week of gestational age)[261]. As we shall now see, there is much more to say about this issue, even if that temporal milestone in foetal development is bound to remain one of the focal points of our practical reasoning.

In the medical literature, the issue of foetal consciousness has received great attention because of the clinical importance of the related issue of foetal pain. For suppose that we conclude that we have sufficient reasons to kill a particular foetus. If on the balance of evidence available we are justified in believing that it is capable of experiencing pain, we would have additional reasons to choose a procedure for killing it, or for terminating the pregnancy, which minimises its suffering or even prevents the foetus form experiencing pain altogether. This

[261] It is important not to confuse embryological or developmental age, computed from the embryo's day of conception, with gestational age, calculated from the first day of the pregnant woman's last menstrual period. Gestational age is roughly two weeks greater than developmental age. Unless stated otherwise, all dates in the coming discussion refer to developmental age. On this distinction see Lee *et al.* (2005, p. 949).

may have an impact on current practices of anaesthesia and analgesia not only during surgical abortions, but also during prenatal interventions on foetuses. Since the experience of pain presupposes consciousness, much of the discussion has centred on determining what structural, functional and environmental conditions correlate with conscious experience and when it is justified to claim that they are satisfied in the foetus.

From a methodological point of view it must be borne in mind that due to the subjective nature of conscious experience and the obvious preverbal nature of foetuses, all evidence of foetal mentation or lack thereof we might have shall necessarily be indirect in character. This warrants a modicum of caution regarding the interpretation of certain kinds of evidence, such as behavioural and physiological responses to external stimuli. For example, it has been observed that by the twelfth week after fertilisation most of the foetus' body responds to touch by moving away[262]. Studies have also shown that by its sixteenth week, when presented with noxious –i.e., potentially damaging– stimuli, the foetus exhibits a stress response similar to that of older children and adults[263]. Should responses of this sort be considered merely indicative of the existence of well-functioning unconscious processes or rather as suggestive of the presence of the conscious experience of pain? As it will become apparent, a reasoned choice of an answer for this question will largely depend on the theory about the correlates of consciousness that one endorses.

Now since cognitive scientists disagree about what is the best such theory, it will result that we shall not have a completely satisfying answer about the onset of the foetal capacity for consciousness. Lacking competence on these issues, I must limit myself to the reconstruction, as I take it to be, of the main competing

[262] See Glover & Fisk (1996).

[263] See Derbyshire (2008, p. 120), Royal College of Obstetricians and Gynaecologists (2010, pp. 14-15, 23). A further kind of foetal response to noxious stimuli which has been observed, this time by its sixteenth week, is the increase of blood flow towards the brain when subjected to venepuncture and transfusions (Glover & Fisk, 1996; Lee, Ralston, Drey, Partridge, & Rosen, 2005, p. 950; Lowery, Hardman, Manning, Hall, & Anand, 2007, p. 279).

general views about the correlates of consciousness. This will occupy the first part of the discussion. Secondly, I shall consider the possibility that the foetus is never truly awake in utero, even granting that it possesses a capacity for consciousness. Finally, I will dwell on the implications these facts may have on our practical reasons.

The majority view regarding the correlates of consciousness is the one that has been aptly called the *corticocentric perspective*[264]. According to this position the capacity for consciousness is entirely based on the cortex. This would allow us to dismiss the presence of a mature stress response to potentially damaging stimuli as indicative of an experience of pain, since these responses are processed by structures located below the cortex[265]. Regarding the interpretation of the perceived behaviour of moving away from aversive stimuli, however, one has to bring to bear the evidence available on the development of the neural pathways that transmit information about external stimuli to the cortex. Thus, for example, if at the time the observed behaviour occurs these pathways have not yet reached the cortex, or if there is good reason to believe that they are not yet functional, then this kind of foetal responses can be deemed to be merely reflex.

On this account of the neural correlates of consciousness, it seems that the earliest time at which we might justifiably expect a foetus to acquire a capacity for conscious experience is around the twenty-first week after fertilisation[266]. This is because it is at this moment in foetal development when fibres directly connecting the cortex with the receptors of internal and external stimuli have

[264] I borrow the terminology from Merker (2007a; 2007b). This is the view assumed throughout the discussions about foetal pain found in the review by Lee *at al.* (2005) published in the *Journal of the American Medical Association* (and endorsed in a 2013 communiqué by the American Congress of Obstetricians and Gynecologists), as well as in the report issued by the Royal College of Obstetricians and Gynaecologists (2010).

[265] This is a point made by many authors, but see, for instance, Mellor *et al.* (2005, p. 457) or Lowery *et al.* (2007, pp. 276, 279-280).

[266] See Lee *et al.* (2005, p. 949), Derbyshire (2006, p. 910; 2008, p. 119).

begun to emerge, so that *some* pathways for the transmission of information can be said to be in place. This does not mean, however, that by this time the biological system of consciousness is either complete or functional. In the case of pain perception, for instance, that would not take place, on an optimistic estimate, until around the twenty-second week[267]. Now, neurons communicate with each other by means of electrical impulses. These impulses are therefore evidence of brain activity, which can be registered by electrodes and subsequently recorded in encephalograms (EEG). It is known that some EEG patterns represent wakefulness. Suppose that, as some authors do[268], we required the presence of these EEG patterns as further evidence of the capacity for conscious experience. It is thus argued that we could not justifiably assert that the foetus gains a mind until its twenty-seventh or twenty-eighth week[269]. Dishearteningly, then, it is not possible to simply identify one week when the foetus acquires a mind. We can, however, on this majority view about consciousness, delineate a period when that occurs, even if one with fuzzy boundaries –for we have reasons to believe that the foetus becomes capable of consciousness no earlier than around its twenty-first week and no later than its twenty-eighth.

Yet, as I announced above, there is a second, emergent view about the neural correlates of consciousness. On this *mesoencephalic perspective*, the functionality of several structures in the brainstem (and, thus, located below the

[267] As estimated by Derbyshire (2010, p. 649). It is interesting to note how new evidence led this author to a re-evaluation of his beliefs, for in previous works he insisted that the biological system for pain experience couldn't be intact and functional before the twenty-fourth week after fertilisation (1996; 2006, p. 910; 2008, p. 120).

[268] Such as Lee *et al.* (2005, pp. 950-952).

[269] Lee *et al.* (2005, p. 950) It must be borne in mind that the presence of the relevant EEG patterns that Lee *et al.* allude to was observed in preterm neonates. Nevertheless, they do not hesitate to extrapolate these findings to foetuses. It has been suggested, however, that such inference might be unwarranted (see below the discussion about the sleep-like state of the foetus in utero). A similar interpretation of the evidence is suggested Lowery *et al.* (2007, p. 279).

cerebral cortex) would be sufficient for a rudimentary form of consciousness[270]. The problem seems to be that there is some evidence that individuals without a cerebral cortex, contrary to what the corticocentric perspective predicts, do have a capacity for conscious experience[271]. As a consequence, adherents to this view suggest that, then, we ought to lower the temporal threshold of possibility for foetal consciousness, proposals ranging from the eighteenth[272] to the twentieth week after fertilisation[273]. Given these discrepancies among the experts, I shall assume a conservative stance, and grant that it is justified to believe that by its eighteenth week of development the foetus acquires a mind.

In parallel with this discussion about the emergence of the *capacity* for consciousness, there is a second, related controversy about when the foetus may first begin to *exercise* it. Indeed, the proposal has been set forth that the foetus is never awake in utero. Even conceding that it becomes capable of conscious experience as early as the most optimistic estimates predict (that is, by its eighteenth week), this proposal would entail that we ought not to expect

[270] This view was lately proposed by Björn Merker (2007a; 2007b).

[271] Thus, Merker (2007a, p. 74) highlights experiments with cats and rodents which, in spite of lacking a cerebral cortex, nevertheless engaged in a series of complex behaviours –such as orientation, exploration, defence, scavenging and eating. Merker suggests that such behaviour is best accounted for by attributing consciousness to the animal. Similarly, he claims to possess behavioural evidence of consciousness in children afflicted with a particular form of anencephaly obtained through his first-hand experience with them and by reports from their families. According to Merker (2007a, p. 79) "[t]hese children are not only awake and often alert, but show responsiveness to their surroundings in the form of emotional or orienting reactions to environmental events [...]. They express pleasure by smiling and laughter, and aversion by "fussing," arching of the back and crying (in many gradations), their faces being animated by these emotional states [...].The children respond differentially to the voice and initiatives of familiars, and show preferences for certain situations and stimuli over others, such as a specific familiar toy, tune, or video program, and apparently can even come to expect their regular presence in the course of recurrent daily routines". Earlier suggestions that the corticocentric perspective cannot account for the behaviour of anencephalic infants can be found in McCullagh (1997) and Saunders (1997).

[272] Brusseau & Mashur (2007, p. 87); Condic (2013, p. 8).

[273] See the report *Fetal Pain: The Evidence* (Doctors on Fetal Pain, 2013).

the foetus to exercise it until after birth. Thus, Rigatto *et al.* observed during their studies on foetal lambs that most of the time their EEG activity correlates with sleep states and that in utero they do not behave as if awake. Yet they noticed that after birth the behavioural changes experienced by the foetus resembled its "awakening" from a "state of profound inhibition"[274]. More recently, Mellor *et al.* suggested that this can be explained by the presence of certain chemical and environmental factors which suppress wakefulness in the prenatal child, keeping it in a sedated state[275]. These sleep-like states of the foetus, alternating from active to quiet sleep, would account for ninety-five per cent of its EEG activity[276]. The other five per cent of such activity could only be considered, on reflection, and according to these authors, an indeterminate state of sleep, but not wakefulness. For one, it is argued, it is too brief and clearly much related to the transition between the sleep episodes that precede and follow it. For another, it is very similar to what is called "post-natal sleep-arousal". This is a state experienced by infants which allows for some cardiorespiratory and postural changes without their becoming awake and without activation of the cortex, thereby preserving the integrity of sleep[277]. As to the profoundness of these sleep-like states in the foetus, the authors further claim that the data demonstrates that threatening stimuli that would normally lead a preterm newborn to full arousal from sleep (such as insufficient oxygen supply or excessive carbon dioxide in the blood) do nothing but further suppress arousal in the foetus[278]. It must be noted that the claim that true wakefulness

[274] Rigatto *et al.* (1986, p. 161).

[275] Mellor *et al.* (2005, pp. 461-464). Such chemical factors include anaesthetics, EEG suppressors and several sleep-inducing agents. Environmental factors include, on the one hand, warmth and, on the other, caused by the amniotic fluid, minimised tactile stimulation as well as, presumably, a reduction of sensations associated with gravity.

[276] Mellor *et al.* (2005, p. 457). Here active sleep refers to rapid-eye-movement (REM) sleep, whereas quiet sleep refers to no-rapid-eye-movement (NREM) sleep. In the foetus, REM sleep is characterised by "breathing, swallowing, licking, and eye movements, and atonia [i.e., absence of muscle tone]", whilst during NREM sleep it exhibits "apnea [i.e., pauses in breathing], absence of eye movements, and tonic muscle activity" (Mellor *et al.*, ibid).

[277] Mellor *et al.* (2005, p. 459).

[278] Mellor *et al.* (2005, p. 460).

only begins after birth is not incompatible with the presence of EEG patterns representing wakefulness by the twenty-seventh or twenty-eighth week of the child's development. As was mentioned above, these were measures of brain activity taken from preterm neonates, not foetuses in utero. Assuming that the views advanced by Rigatto *et al.* and Mellor *et al.* are true, this is precisely one difference we ought to expect to exist between foetuses and sufficiently developed preterm neonates[279].

What does this all entail? Let us reflect first on the issue of the acquisition of the capacity for consciousness. First, ontology. As I explain in CH. III, **§2**, some Lockean views posit more stringent requirements than others in their accounts of what we are essentially. Consider what is conceivably the least demanding of such accounts, in which merely the presence of a capacity for consciousness is required, without the need for mental content. Even on this view none of us could have begun to exist before the eighteenth week after conception, for on a conservative estimate this is the earliest time at which it can justifiably be claimed that the foetus acquires a mind.

Second, practical rationality. Here I assume, as in CH.III, **§3** I argued that we ought to, that the relation with our future that matters is psychological in nature. Again, let us suppose that we embrace a minimalist account of such a relation. We may believe, as Jeff McMahan does, that continuity of the capacity for consciousness is a sufficient condition for the relation that matters to obtain. Because of the scientific evidence just mentioned, we could justifiably believe that by its eighteenth week of development the foetus' future life is personally

[279] What credence does this view merit? Derbyshire uses Rigatto *et al.* (1986) and Mellor *et al.* (2005) to defend that it is "seemingly reasonable" to assert that "pain [is] objectively impossible in utero" (2010, p. 650). And, for their part, the authors of the RCOG report are more cautious in their conclusions (2010, p. 23). Certainly they admit that "increasing evidence" supports the view that the foetus is never awake in utero, yet they immediately emphasise that "this cannot be known with certainty". They do, however, extract a more straightforward methodological caveat regarding "the potential pitfalls of extrapolating observations in newborn preterm infants to a fetus of the same gestational age".

valuable for him. But, as the reader may remember, on McMahan's account the mere obtainment of sameness of the capacity for consciousness, in the absence of continuity of mental content, implies that the relation that matters holds only at the minimum possible degree. Therefore, the value of the events that constitute the life of a being for which this is true must be heavily discounted. Regarding the foetus, this implies that the time-relative personal value its future life now has is considered exceedingly low. Perhaps it is even more so than we used to believe. For on the traditional view which assumed the wakefulness of the foetus in utero, it was reasonable to assume that as structural changes augmented the sophistication of its capacity for consciousness, the foetus actually developed sharper experiences of its environment, consequently forming fitting belief-like and desire-like states. This allows for the claim that over time the foetus becomes more intensely psychologically connected with its future, even if marginally so, as the pregnancy advances. In this way, we could accommodate the widespread intuition that our reasons against killing a foetus, as given by the personal value of its life, grow stronger as birth approaches. But, of course, this does not follow if we assume that this traditional view is wrong and that in fact human beings never experience true wakefulness until after birth. On this alternative view it is doubtful that the foetus's psychological isolation from its future ameliorates over time as much as we were wont to believe, for it remains in a sleep-like state – and possibly a dreamless sleep at that[280].

We must, then, distinguish between two different questions about the onset foetal consciousness –one regarding the acquisition of the capacity for conscious experience, the other regarding the possibility of its exercise. As to the question about the foetus' acquisition of said capacity, I have attempted to show the differences between a majority view about the correlates of consciousness, which makes it depend entirely on the cortex, and a minority view, which claims that several subcortical structures are sufficient for a simple form of conscious experience. Assuming that it is better to err on the side of

[280] Mellor *et al.* (2005, p. 461)

caution, I shall argue throughout the main text of the thesis as if the minority view was true. This is because it is the one which defends the earliest start of the capacity for consciousness, placing it around the eighteenth week after conception. As to whether the exercise of this capacity is somehow inhibited in utero, this is also a disputed matter. Certainly, we ought to contemplate the possibility that the foetus is never awake while inside the womb and, if this proved to be true, to modify our attitudes accordingly. Reasoning on the same principle of caution, however, I shall argue as if that this view was false. Consequently, I will assume that the foetus' conscious experiences increase in sophistication as the pregnancy advances and the relevant biological structures, whatever those might be, mature.

Bibliography

Abbott, P. (1978). Philosophers and the Abortion Question. *Political Theory , 6* (3), 313-335.

American Congress of Obstetricians and Gynecologists. (2013, June 20). Facts Are Important: Fetal Pain.

Baker, L. R. (2002b). Brief Reply to Rosenkrantz's Comments on My "The Ontological Status of Persons". *Philosophy and Phenomenological Research , 65* (2), 394-396.

Baker, L. R. (2002a). The Ontological Status of Persons. *Philosophy and Phenomenological Research , 65* (2), 370-388.

Baker, L. R. (1999a). Unity without Identity: A New Look at Material Constitution. *Midwest Studies in Philosophy , XXIII*, 144-165.

Baker, L. R. (1999b). What Am I? *Philosophy and Phenomenological Research , 59* (1), 151-159.

Baker, L. R. (1997). Why Constitution is Not Identity. *The Journal of Philosophy , 94* (12), 599-621.

Bassen, P. (1982). Present Stakes and Future Prospects: The Status of Early Abortions. *Philosophy and Public Affairs , 11* (4), 314-337.

Becker, L. C. (1975). Human Being: The Boundaries of the Concept. *Philosophy and Public Affairs , 4* (4), 334-359.

Boonin, D. (2003). *A Defense of Abortion.* Cambridge: Cambridge University Press.

Brown, M. T. (2002). A Future Like Ours Revisited. *Journal of Medical Ethics , 28*, 192-195.

Brusseau, R. R., & Mashour, G. A. (2007). Subcortical consciousness: Implications for fetal anesthesia and analgesia. *Behavioral and Brain Sciences , 30* (1), 86-87.

Buss, M. J. (1967). The Beginning of Human Life as an Ethical Problem. *The Journal of Religion , 47* (3), 244-255.

Callahan, D. (1970). *Abortion: Law, Choice and Morality.* London: The Macmillan Company.

Carter, W. R. (1982). Do Zygotes Become People? *Mind , New Series, 91* (361), 77-95.

Catholic Church. (1993). *Catechism of the Catholic Church.* Retrieved from Vatican: The Holy See: http://www.vatican.va/archive/ENG0015/_INDEX.HTM

Catholic Church. (1995, March 25). *Evangelium Vitae.* Retrieved from Vatican: The Holy See: http://www.vatican.va/holy_father/john_paul_ii/encyclicals/documents/hf_jp-ii_enc_25031995_evangelium-vitae_en.html

Catholic Church. (1965, December 7). *Pastoral Constitution of the Church in the Modern World - Gaudium et Spes.* Retrieved September 18, 2012, from Vatican: The Holy See: http://www.vatican.va/archive/hist_councils/ii_vatican_council/documents/vat-ii_const_19651207_gaudium-et-spes_en.html

Condic, M. L. (2013, May 23). Testimony of Maureen L. Condic, Ph.D. University of Utah, School of Medicine, Department of Neurobiology and Anatomy Before the Subcommittee on the Constitution and Civil Justice, Committee on the Judiciary, U.S. House of Representatives.

Cudd, A. E. (1990). Sensationalized Philosophy: A Reply to Marquis's "Why Abortion is Immoral". *The Journal of Philosophy , 87* (5), 262-264.

DeGrazia, D. (2012). *Creation Ethics. Reproduction, Genetics and Quality of Life.* Oxford: Oxford University Press.

DeGrazia, D. (2005). *Human Identity and Bioethics.* New York: Cambridge University Press.

Denton, M. (1986). *Evolution: A Theory in Crisis.* Bethesda: Adler & Adler.

Derbyshire, S. W. (1997). Analgesic and anaesthetic procedures are being introduced because of shoddy sentimental argument. *British Medical Journal , 314*, 1201.

Derbyshire, S. W. (2006). Can fetuses feel pain? *British Medical Journal , 332*, 909-912.

Derbyshire, S. W. (2008). Fetal Pain: Do We Know Enough to Do the Right Thing? *Reproductive Health Matters , 16* (31 Supplement), 117-126.

Derbyshire, S. W. (2010). Foetal pain? *Best Practice & Research Clinical Obstetrics and Gynaecology , 24* (5), 647-655.

Derbyshire, S. W., & Furedi, A. (1996). "Fetal pain" is a misnomer. *Bristish Medical Journal , 313*, 795.

Devine, P. E. (1984). The Scope of the Prohibition Against Killing. In J. Feinberg, *The Problem of Abortion* (pp. 21-42). Belmont: Wadsworth Publishing Company.

Diamond, J. J. (1975). Abortion, animation, and biological hominization. *Theological Studies* (36), 305-324.

Doctors on Fetal Pain. (2013). *Fetal Pain: The Evidence.* www.doctorsonfetalpain.org.

Donagan, A. (1985). Comments on Dan Brock and Terrence Reynolds. *Ethics , 95* (4), 874-886.

Dworkin, R. (1994). *Life's Dominion. An Argument about Abortion, Euthanasia and Individual Freedom.* New York: Vintage Books.

Dworkin, R. (2006). Rawls and the Law. In R. Dworkin, *Justice in Robes* (pp. 241-295). Cambridge: Harvard University Press.

Engelhardt, H. T. (1974). The Ontology of Abortion. *Ethics , 84* (3), 217-234.

English, J. (1984). Abortion and the Concept of a Person. In J. Feinberg (Ed.), *The Problem of Abortion* (Second ed., pp. 151-160). Belmont: Wadsworth Publishing Company.

Feinberg, J. (Ed.). (1984). *The Problem of Abortion* (Second ed.). Belmont: Wadsworth Publishing Company.

Finnis, J. (2011d). Euthanasia and Justice. In J. Finnis, *Human Rights and Common Good: Collected Essays* (Oxford Scholarship Online ed., Vol. III, pp. 212-239). Oxford: Oxford University Press.

Finnis, J. (2011c). Intentionally Killing the 'Permanently Unconscious'. In J. Finnis, *Intention and Identity: Collected Essays* (Oxford Scholarship Online ed., Vol. II, pp. 313-322). Oxford: Oxford University Press.

Finnis, J. (2011e). Justice for Mother and Child. In J. Finnis, *Human Rights and Common Good: Collective Essays* (Oxford Scholarship Online ed., Vol. III, pp. 307-315). Oxford: Oxoford University Press.

Finnis, J. (2011f). Nuclear Deterrence and the End of Christendom. In *Religion and Public Reasons: Collected Essays* (Oxford Scholarship Online ed., Vol. V, pp. 275-289). Oxford: Oxford University Press.

Finnis, J. (2011a). Organic Unity, Brain Life, and Our Beginning. In J. Finnis, *Intention and Identity: Collected Essays* (Oxford Scholarship Online ed., Vol. II, pp. 279-286). Oxford: Oxford University Press.

Finnis, J. (2011g). The Consistent 'Ethic of Life'. In J. Finnis, *Religion and Public Reasons: Colllected Essays* (Oxford Scholarship Online ed., Vol. V, pp. 291-328). Oxford: Oxford University Press.

Finnis, J. (1973). The Rights and Wrongs of Abortion: A Reply to Judith Thomson. *Philosophy and Public Affairs , 2* (2), 117-145.

Finnis, J. (2011b). When Most People Begin. In J. Finnis, *Intention and Identity: Collected Essays* (Oxford Scholarship Online ed., Vol. II, pp. 287-291). Oxford: Oxford University Press.

Freeman, S. (2004). Public Reason and Political Justifications. *Fordham Law Review , 72* (5), 2021-2072.

George, R. P. (2013). Infanticide and madness. *Journal of Medical Ethics , 39* (5), 299-301.

Gerber, D. (1972). Abortion: The Uptake Argument. *Ethics , 83* (1), 80-83.

Gerber, R. J. (1972). Abortion: Paramaters for Decision. *Ethics , 82* (2), 134-154.

Gillespie, N. C. (1977). Abortion and Human Rights. *Ethics , 87* (3), 247-243.

Giubilini, A., & Minerva, F. (2012). After-birth abortion: why should the baby live? *Journal of Medical Ethics* (published online February 23, 2012).

Glover, J. (1977). *Causing Death and Saving Lives* (1990 ed.). London: Penguin Books.

Glover, V., & Fisk, N. (1996). We don't know; better to err on the safe side from mid-gestation. *British Medical Journal , 313*, 796.

Grice, H. P. (1941). Personal Identity. *Mind , 50* (200), 330-350.

Hardin, G. (1968). Abortion, Or Compulsory Pregnancy? *Journal of Marriage and Family , 30* (2), 246-251.

Hare, R. M. (1975). Abortion and the Golden Rule. *Philisophy and Public Affairs , 4* (4), 201-222.

Harman, E. (1999). Creation Ethics: The Moral Status of Early Fetuses and the Ethics of Abortion. *Philosophy and Public Affairs , 28* (4), 310-324.

Harris, J. (1985). *The Value of Life.* London: Routledge.

Hodge, H. L. (1869). *Foeticide, Or Criminal Abortion. A Lecture Introductory to the Course on Obstetrics, and Diseases of Women and Children. University of Pennsylvania. Session 1839-1840.* Lindsay and Blakiston.

Holtug, N. (2010). *Persons, Interests, and Justice* (Oxford Scholarship Online ed.). Oxford: Oxford University Press.

Hugdahl, K. (2005). Symmetry and Assymmetry in the Human Brain. *European Review , 13* (2), 119-133.

Hull, D. L. (1998). On Human Nature. In D. L. Hull, & M. Ruse, *The Philosophy of Biology* (pp. 383-397). New York: Oxford University Press.

Johnston, M. (1987). Human Beings. *The Journal of Philosophy , 84* (2), 59-83.

Kaczor, C. (2011). *The Ethics of Abortion: Women's Rights, Human Life and the Question of Justice.* London: Routledge.

Kagan, S. (2012). *Death.* New Haven: Yale University Press.

Kamm, F. M. (2007). *Intrincate Ethics.* New York: Oxford University Press.

Keown, J. (2012). *The Law and Ethics of Medicine. Essays on the Inviolability of Human Life.* Oxford: Oxford University Press.

Korcz, K. A. (2004). Reply to Marquis's 'Kocz's Objections to the Future-of-Value Argument'. *Journal of Social Philosophy , 35* (1), 61-65.

Korcz, K. A. (2002). Two Moral Strategies Regarding Abortion. *Journal of Social Philosophy , 33* (4), 581-565.

Kuhse, H., & Singer, P. (1985). *Should the Baby Live? The Problem of Handicapped Infants.* Oxford: Oxford University Press.

Lee, P. (2010). *Abortion & Unborn Human Life* (Second ed.). Washington, D.C.: The Catholic University of America Press.

Lee, S. J., Ralston, H. J., Drey, E. A., Partridge, J. C., & Rosen, M. A. (2005). Fetal Pain. A Systematic Multidisciplinary Review of the Evidence. *Journal of the American Medical Association , 294* (8), 947-954.

Lewis, D. (1983). Survival and Identity. In D. Lewis, *Philosophical Papers* (Oxford Scholarship Online ed., Vol. I, pp. 55-72). Oxford: Oxford University Press.

Locke, J. (2008). Of Identity and Diversity. In J. Perry (Ed.), *Personal Identity* (Second ed., pp. 33-52). London: University of California Press.

Loux, M. J. (2006). *Metaphysics. A Contemporary Introduction* (3rd ed.). New York: Routledge.

Loux, M. J., & Zimmerman, D. W. (Eds.). (2003). *The Oxford Handbook of Metaphysics.* Oxford: Oxford University Press.

Lowe, E. J. (2003). Individuation. In M. J. Loux, & D. W. Zimmerman, *The Oxford Handbook of Metaphysics* (pp. 75-95). Oxford: Oxford University Press.

Lowery, C. L., Hardman, M. P., Manning, N., Hall, R. W., & Anand, K. J. (2007). Neurodevelopmental Changes of Fetal Pain. *Seminars in Perinatology , 7* (4), 275-282.

Margolis, J. (1973). Abortion. *Ethics , 84* (1), 51-61.

Marquis, D. (2002). A defence of the potential future of value theory. *Journal of Medical Ethics* (28), 198-201.

Marquis, D. (2006). Abortion and the Beginning and End of Human Life. *Journal of Law Medicine and Ethics , 34* (1), 16-25.

Marquis, D. (2007a). Abortion Revisited. In B. Steinbock (Ed.), *The Oxford Handbook of Bioethics* (pp. 395-415). Oxford: Oxford University Press.

Marquis, D. (2001). Deprivations, Futures and the Wrongness of Killing. *Journal of Medical Ethics , 27*, 363-369.

Marquis, D. (2004). Korcz's Objections to the Future-of-Value Argument. *Journal of Social Philosophy , 35* (1), 56-65.

Marquis, D. (1998). Reiman on Abortion. *Journal of Social Philosophy , 29* (1), 143-145.

Marquis, D. (2003). Review of Jeff McMahan, 'The Ethics of Killing: Problems at the Margins of Life'. *Ethics , 113* (2), 437-440.

Marquis, D. (2003). Review of McMahan, Jeff. "The Ethics of Killing: Problems at the Margins of Life". *Ethics , 113* (2), 437-440.

Marquis, D. (2005). Savulescu's objections to the Future-of-Value Argument. *Journal of Medical Ethics , 35*, 119-122.

Marquis, D. (2007b). The Moral-Principle Objection to Human Embryonic Stem-Cell Research. *Metaphilosophy , 38* (2-3).

Marquis, D. (1989). Why Abortion is Immoral. *The Journal of Philosophy , 86* (4), 183-202.

McCloskey, H. J. (1975). The Right to Life. *Mind , New Series, 84* (335), 403-425.

McCullagh, P. (1997). Can fetal suffering be excluded beyond reasonable doubt? *British Medical Journal , 314*, 302-303.

McInerney, P. K. (1990). Does a Fetus Already Have a Future-Like-Ours? *The Journal of Philosophy , 87* (5), 264-268.

McMahan, J. (2008). Challenges to Human Equality. *The Journal of Ethics , 12*, 81-104.

McMahan, J. (2007). *Infanticide.* Retrieved June 12, 2013, from http://www.fas.rutgers.edu/cms/phil/dmdocuments/Infanticide.pdf

McMahan, J. (2013). Infanticide and moral consistency. *Journal of Medical Ethics* (39), 273-280.

McMahan, J. (2006). Paradoxes of Abortion and Prenatal Injury. *Ethics , 116* (4), 625-655.

McMahan, J. (2002). *The Ethics of Killing. Problems at the Margins of Life.* Oxford: Oxford University Press.

Mellor, D. J., Diesch, T. J., Gunn, A. J., & Bennet, L. (2005). The importance of 'awareness' for understanding fetal pain. *Brain Research Reviews* (49), 455-471.

Merker, B. (2007b). Author's Response. Grounding consciousness: The mesodiencephalon as thalamocortical base. *Behavioral and Brain Sciences , 30* (1), 110-120.

Merker, B. (2007a). Consciousness without a cerebral cortex: A challenge for neuroscience and medicine. *Behavioral and Brain Sciences , 30* (1), 63-81.

Meyers, C. (2010). *The Fetal Position.* Amherst: Prometheus Books.

Miller-Keane Encyclopedia and Dictionary of Medicine, Nursing, and Allied Health. (2003). *abortion*. Retrieved 10 1, 2012, from The Free Online Medical Dictionary: http://medical-dictionary.thefreedictionary.com/abortion

Mosby's Medical Dictonary. (2009). *abortion*. Retrieved 10 1, 2012, from The Free Online Medical Dictionary: http://medical-dictionary.thefreedictionary.com/abortion

Nagel, T. (1970). *The Possibility of Altruism.* Oxford: Oxford University Press.

Newton, L. (1975). Humans and Persons: A Reply to Tristram Engelhardt. *Ethics , 85* (4), 332-336.

Noonan, H. W. (2003). *Personal Identity* (Second ed.). London: Routledge.

Norcross, A. (1990). Killing, Abortion and Contraception: A Reply to Marquis. *The Journal of Philosophy , 87* (5), 268-277.

Nozick, R. (1981). *Philosophical Explanations.* Cambridge: Belknap Press of Harvard University Press.

Oderberg, D. S. (1997). Modal Properties, Moral Status and Identity. *Philosophy and Public Affairs , 26* (3), 259-298.

Olson, E. T. (2010, October 28). *Personal Identity*. Retrieved May 9, 2013, from Stanford Encyclopedia of Philosophy: http://plato.stanford.edu/entries/identity-personal/

Olson, E. T. (1997b). *The Human Animal. Personal Identity without Psychology.* Oxford: Oxford University Press.

Olson, E. T. (1997a). Was I Ever a Fetus? *Philosophy and Phenomenological Research , 57* (1), 95-110.

Olson, E. T. (2007). *What Are We? A Study in Personal Ontology.* New York: Oxford University Press.

Parfit, D. (1982). Future Generations: Further Problems. *Philosophy and Public Affairs* (11), 113-172.

Parfit, D. (2011a). *On What Matters* (Vol. I). Oxford: Oxford University Press.

Parfit, D. (2011b). *On What Matters* (Vol. II). Oxford: Oxford University Press.

Parfit, D. (1971). Personal Identity. *The Philosophical Review , 80* (1), 3-27.

Parfit, D. (2008). Persons, Bodies and Human Beings. In T. Sider, J. Hawthorne, & D. W. Zimmerman, *Contemporary Debates in Metaphysics* (pp. 177-208). Oxford: Blackwell Publishing.

Parfit, D. (1984). *Reasons and Persons* (1987 rev. ed.). Oxford: Oxford University Press.

Parfit, D. (1995). The Unimportance of Identity. In H. Harris (Ed.), *Identity* (pp. 13-45). Oxford: Clarendon Press.

Parfit, D. (2012). We Are Not Human Beings. *Philosophy* (87), 5-28.

Parsons, S. J. (2002). Present Self-represented Futures of Value are a Reason for the Wrongness of Killing. *Journal of Medical Ethics , 28*, 196-197.

Perry, J. (Ed.). (2008). *Personal Identity* (Second ed.). London: University of California Press.

Perry, J. (2008 [1975]). Personal Identity, Memory, and the Problem of Circularity. In J. Perry (Ed.), *Personal Identity* (Second ed., pp. 135-155). London: University of California Press.

Perry, J. (2008a). The Problem of Personal Identity. In J. Perry (Ed.), *Personal Identity* (Second ed., pp. 3-30). London: University of California Press.

Perry, M. J. (2007). *Toward a Theory of Human Rights: Religion, Law, Courts.* New York: Cambridge University Press.

Pluhar, W. S. (1977). Abortion and Simple Consciousness. *The Journal of Philosophy , 74* (3), 159-172.

Quinn, W. (1984). Abortion, Identity and Loss. *Philosophy and Public Affairs , 13* (1), 24-54.

Rawls, J. (1999 [1971]). *A Theory of Justice* (revised ed.). Cambridge: Harvard University Press.

Rawls, J. (2005 [1993]). *Political Liberalism. Expanded Edition.* New York: Columbia University Press.

Raz, J. (2002). *Engaging Reason: On the Theory of Value and Action* (Oxford Scholarship Online ed.). Oxford: Oxford University Press.

Raz, J. (1975). *Practical Reason and Norms.* Princeton: Princeton University Press.

Reiman, J. (1996). Abortion, Infanticide, and the Asymmetric Value of Human Life. *Journal of Social Philosophy , 27* (3), 181-200.

Riddle, J. M. (1992). *Contraception and Abortion from the Ancient World to the Renaissance.* Cambridge: Harvard University Press.

Riddle, J. M. (1997). *Eve's Herbs. A History of Contraception and Abortion in the West.* Cambrigde: Harvard University Press.

Rigatto, H., Moore, M., & Cates, D. (1986). Fetal breathing and behavior measured through a double-wall Plexiglas window in sheep. *Journal of Applied Physiology , 61*, 160-164.

Rokyta, R. (2008). Fetal Pain. *International Journal of Prenatal and Perinatal Psychology and Medicine , 20* (3/4), 167-178.

Ross, S. L. (1982). Abortion and the Death of the Fetus. *Philosophy and Public Affairs , 11* (3), 232-245.

Royal College of Obstetricians and Gynaecologists. (2010). *Fetal Awareness. Review of Research and Recommendations for Parctice. Report of a Working Party.* London: Royal College of Obstetricians and Gynaecologists.

Sadler, T. W. (2012). *Langman's Medical Emrbyology* (Twelfth ed.). Baltimore: Lippincott Williams & Wilkins.

Saunders, P. J. (1997). We should give them the benefit of the doubt. *British Medical Journal , 314*, 303.

Savulescu, J. (2002). Abortion, Embryo Destruction and the Future of Value Argument. *Journal of Medical Ethics , 28*, 133-125.

Selgelid, M. J. (2012). Moral uncertainty and the moral status of early human life. *Monash Bioethics Review , 30* (1), 52-57.

Sher, G. (1977). Hare, Abortion and the Golden Rule. *Philosophy and Public Affairs , 6* (2), 185-190.

Shoemaker, S. (1984). Personal Identity: a Materialist's Account. In S. Shoemaker, & R. Swinburne, *Personal Identity* (pp. 67-132). Oxford: Basil Blackwell.

Shoemaker, S. (1970). Persons and Their Pasts. *American Philosophical Quarterly , 7* (4), 269-285.

Shoemaker, S. (1997). Self and Substance. *Noûs , 31* (Supplement: Philosophical Perspectives, 11, Mind, Causation, and World), 283-304.

Shoemaker, S. (1963). *Self-Knowledge and Self-Identity.* New York: Cornell University Press.

Shoemaker, S. (1992). Unger's Psychological Continuity Theory. *Philosophy and Phenomenological Research , 52* (1), 139-143.

Singer, P. (2011). *Practical Ethics* (3rd ed.). New York: Cambridge University Press.

Singer, P. (1995). *Rethinking Life and Death. The Collapse of Our Traditional Ethics.* Oxford: Oxford University Press.

Singer, P. (2011b, March 4). *Taking Life: Animals. On the Human - comment-4860*. Retrieved June 1, 2013, from On the Human: http://onthehuman.org/2011/02/taking-life-animals/comment-page-1/#comment-4860

Snowdon, P. (1990). Persons, animals and ourselves. In C. Gill (Ed.), *The Person and the Human Mind: Issues in Ancient and Modern Philosophy.* Oxford: Oxford University Press.

Steadman's Medical Dictionary. (2006). *Abortion - Medical Definition*. Retrieved 10 1, 2012, from Medilexicon: http://www.medilexicon.com/medicaldictionary.php?t=143

Steinbock, B. (2011). *Life Before Birth. The Moral and Legal Status of Embryos and Fetuses* (2nd ed.). Oxford: Oxford University Press.

Steinbock, B. (1992). *Life Before Birth: The Moral and Legal Status of Embryos and Fetuses* (1st ed.). Oxford: Oxford University Press.

Steinbock, B. (2007). Moral Status, Moral Value and Human Embryos: Implications for Stem Cell Research. In B. Steinbock (Ed.), *The Oxford Handbook of Bioethics* (pp. 416-430). Oxford: Oxford University Press.

Steinbock, B. (2006). The Morality of Killing Human Embryos. *Journal of Law, Medicine and Ethics , 34* (1), 26-34.

Strawson, P. F. (1959). *Individuals. An Essay in Descriptive Metaphysics.* London: Routledge.

Thomson, J. J. (1971). A Defense of Abortion. *Philosophy and Public Affairs , 1* (1), 47-66.

Thomson, J. J. (1997). People and their Bodies. In J. Dancy (Ed.), *Reading Parfit* (pp. 202-229). Oxford: Blackwell Publishing.

Tighe, M. (2006). Fetuses can feel pain. *British Medical Journal , 332*, 1036.

Tooley, M. (1972). Abortion and Infanticide. *Philosophy and Public Affairs , 2* (1), 47-66.

Tooley, M. (1983). *Abortion and Infanticide.* New York: Oxford University Press.

Tooley, M. (2009). Abortion: Why a Liberal View Is Correct. In M. Tooley, C. Wolf-Devine, P. E. Devine, & A. M. Jaggar, *Abortion: Three Perspectives* (pp. 3-64). New York: Oxford University Press.

Tooley, M., Wolf-Devine, C., Devine, P. E., & Jaggar, A. M. (2009). *Abortion: Three Perspectives.* New York: Oxford University Press.

Trammell, R. L. (1976). Tooley's Moral Symmetry Principle. *Philosophy and Public Affairs , 5* (3), 305-313.

Trussell, J., & Raymond, E. G. (2012, September). *Emergency Contraception: Al Last Chance to Prevent Unintended Pregnancy.* Retrieved November 29, 2012, from Emergency Contraception: http://ec.princeton.edu/questions/ec-review.pdf

Unger, P. (1990). *Identity, Consciousness and Value.* New York: Oxford University Press.

Unger, P. (2000). The Survival of the Sentient. *Noûs , 34* (Supplement: Philosophical Perspectives, 14, Action and Freedom), 325-348.

van Inwagen, P. (1990). *Material Beings.* Ithaca: Cornell University Press.

Warren, M. A. (1997). *Moral Status: Obligations to Persons and Other Living Things.* New York: Oxford University Press.

Warren, M. A. (1984). On the Moral and Legal Status of Abortion (with a Postscript on Infanticide). In J. Feinberg, *The Problem of Abortion* (2nd ed., pp. 102-119). Belmont: Wadsworth.

Warren, M. A. (1973). On the Moral and Legal Status of Abortion. *The Monist , 57* (4).

Wertheimer, R. (1971). Understanding the Abortion Argument. *Philosophy and Public Affairs , 1* (1), 67-95.

Wiggins, D. (2001). *Sameness and Substance Renewed.* Cambridge: Cambridge University Press.

Williams, B. (1973a). Personal Identity and Individuation. In B. Williams, *Problems of the Self. Philosophical Papers 1956-1972* (pp. 1-18). Cambridge: Cambridge University Press.

Williams, B. (1973b). The Self and the Future. In B. Williams, *Problems of the Self. Philosophical Papers 1956-1972* (pp. 46-63). Cambridge: Cambridge University Press.

World Health Organization. (2012). *Safe Abortion: technical and policy guidance for health systems. Second Edition.* Geneva: World Health Organization.

Zaitchik, A. (1981). Viability and the Morality of Abortion. *Philosophy and Public Affairs , 10* (1), 18-26.

Zimmerman, D. W. (2003). Material People. In M. J. Loux, & D. W. Zimmerman (Eds.), *The Oxford Handbook of Metaphysics* (pp. 491-526). Oxford: Oxford University Press.