Etica & Politica / Ethics & Politics, X, 2008, 2, pp. 72-108

Prudence and Morality in Butler, Sidgwick, and Parfit

Alessio Vaccari
Università di Roma "La Sapienza"
alessio.vaccari@gmail.com

ABSTRACT
The debate on personal identity has profoundly modified the approach to the analysis of prudence, its structure and its links with rationality and morality. While in ethics of 18th and 19th centuries the problem of justifying prudent behaviour rationally did not exist, in contemporary ethics it seems no longer possible to justify it rationally. Particularly, from the perspective of the complex account of personal identity it seems that the only way to condemn great imprudence is from the point of view of morality. In this way we assist to a slow erosion of the clear-cut distinction between prudence and morality. The paper illustrates this change contrasting the analysis of prudence made by Joseph Butler, and then followed by his heir Henry Sidgwick, with that recently made by Derek Parfit.

1. *Introduction*

The recent debate on personal identity has profoundly modified the approach to the analysis of prudence, its structure and its links with altruism and moral theory. In contemporary thinking we witness a slow erosion of the clear-cut categorical distinction between egoism and altruism characteristic of the philosophical framework of the English 18th century[1].

In 18th century ethics, in which the problem of rationally justifying prudent behaviour did not exist, a particularly urgent need was felt to find a solution to the issue involving the possibility of accepting prudence from the specifically moral standpoint. If prudence is to be considered as a completely self-interested line of behaviour and, according to the thesis prevailing at the time, the moral point of view may be identified with an impartial and disinterested outlook, the question naturally arose of whether prudence could be reconciled, and in what way, with the need for morally virtuous behaviour.

With reference to this problem, the project of Shaftesbury and Hutcheson, in which it was claimed that a comparatively clear-cut difference existed between prudence and moral virtue, was opposed by the con-

[1] For a reconstruction of the relationship between prudence and ethics within the framework of 18th century English philosophy and the changes wrought in contemporary thinking see Eugenio Lecaldano, *L'etica e l'identità personale: tra prudenza e azione razionale*, *Archivio di filosofia*, LV n. 1-3-, pp. 231-259.

ciliatory proposal of Butler, and Smith, in which prudence was deemed to be a behaviour that did not clash with moral virtue, without however coinciding with it.

For our purpose it is important to point out that, even those who were inclined to believe a relatively strong degree of reconciliation was possible between ethical behaviour and prudence accepted the general thesis that it was possible to make a categorical distinction between these two levels of behaviour: egoistic actions motivated by self-interest ought not to be confused with altruistic or moral actions guided by benevolence.

The distinction between egoism and altruism, and consequently between prudence and ethics, albeit in a radically different philosophical context to that of utilitarianism, is not discussed in the English tradition in a book written in the late nineteenth century such as *Methods of Ethics*[2] by H. Sidgwick. And it is precisely this distinction that leads to what Sidgwick considers the most difficult problem facing ethics: the dualism of the practical reason, that is, the serious and profound contradiction between two ethical principles, that of rational egoism and that of rational benevolence, due to the simultaneous validity of two rational normative intuitions of equal weight and strength.

In contemporary philosophy we witness a shift in the axis of thinking regarding the relations between prudence and ethics versus both the 18th century framework and the theses expressed in the *Methods*. In addition to the problem of the rationality of ethics, the need is felt to raise the increasingly fundamental question of the rationality of prudence.

The turning point is represented by the huge success encountered in recent years by several ideas contained in the treatment of prudence and rational egoism given by Sidgwick in his *Methods*. In one argument, known in contemporary philosophy as the "parity argument"[3], Sidgwick describes the difficulty of considering the point of view of rational egoism fully justified and evident as compared with that of altruism. He writes:

From the point of view, indeed, of abstract philosophy, I do not see why the Egoist principle should pass unchallenged any more than the Univeralistic. I do not see why the axiom of Prudence should not be questioned, when it conflicts with present inclination, on a ground similar to that on which Egoists refuse to admit the axiom of Rational Benevolence. If the Utilitarian has to answer to the question, 'Why should I sacrifice my own happiness for

[2] Henry Sidgwick, *The Methods of Ethics* (1874, 7th ed. 1907), Indianapolis, Ind., Hackett, 1981.

[3] This term was introduced by D.O. Brink in *Sidgwick and the Rationale for Rational Egoism*, in B. Schultz (ed.), *Essays on Henry Sidgwick*, New York, Cambridge University Press, 1991, pp. 199-239.

the greater happiness of another?' it must surely be admissible to ask the Egoist, 'Why should I sacrifice a present pleasure for a greater one in the future? Why should I concern myself about my own future feelings any more than about the feelings of other persons?' It undoubtedly seems to Common Sense paradoxical to ask for a reason why one should seek one's own happiness on the whole; but I do not see how the demand can be repudiated as absurd by those who adopt the views of the estreme empirical school of psychologists, although those views are commonly supposed to have a close affinity with Egoistic Hedonism. Grant that the Ego is merely a system of coherent phenomena, that the permanent identical 'I' is not a fact but a fiction, as Hume and his followers maintain; why, then, should one part of the series of feelings into which the Ego is resolved be concerned with another part of the same series, any more than with any other series?[4]

In this well-known passage, Sidgwick emphasizes the need to provide arguments in support of prudence, and explained how the demand for these arguments derived from an analysis of the structure of prudence involving both its dimension of temporal neutrality and the implications of a complex or atomistic conception of the self. Sidgwick's approach, recently taken up again by authors like T. Nagel[5], R.M. Hare[6] and D. Parfit[7] by means of a further scrutiny of the conditions of prudent action, led to a reappraisal of the general question of the relationship between ethics and prudence.

The most significant results within this new analytical paradigm have been achieved by Derek Parfit. Following in Sidgwick's footsteps he reconstructs prudence as a theory of individual rationality, which he calls "Self-interest Theory" or S. S theory states that each agent must maximize his overall happiness, taking into consideration the probable total duration of his own life. From the content of this substantive objective it implicitly follows that each type of temporal preference must be considered irrational as the agents are asked to have an equal interest in all parts of their lives. Parfit constructs two groups of objections to the self-interest theory, reaching the conclusion that the substantive objective of this theory requires the agents to adopt an attitude vis-à-vis their own future that has no rational justification: the theory must therefore be rejected.

The outcome of these arguments thus shows that, in accordance with the classical theory of prudence, we can no longer consider imprudent actions as

---

[4] Henry Sidgwick, *The Methods of Ethics*, pp. 418-19.

[5] Thomas Nagel, *The Possibility of Altruism,* Princeton, Princeton University Press, 1970.

[6] R. M. Hare, *Moral Thinking. Its Levels, Methods and Point*, New York, Oxford University Press, 1981.

[7] Derek Parfit, *Reasons and Persons*, Oxford, Clarendon Press, 1984.

irrational since prudence does not prescribe rational actions. We therefore need a new theory that, by adopting a different criterion of rationality, will allow us to condemn imprudent actions.

Parfit believes that the only available strategy is to modify our ethical theory in such a way as to extend its application also to the class of imprudent actions which were conventionally not subject to moral evaluation.

This article thus focuses on the examination of the outcomes of the main arguments developed by Parfit to refute the "Self-interest Theory".

Furthermore, it will be attempted to demonstrate why the only valid argument among those presented in *Reasons and Persons*, is the one based on a revised conception of personal identity, thereby confirming our idea to consider Parfit's work as an essential part of the new analytical paradigm that has brought about a change in the relationship between prudence and ethics.

Before analysing the structure of the theory of self-interest and its peculiar features as a theory of rationality, I should like to examine briefly several classical treatments of prudence in the English philosophical thinking. In particular, I shall take into consideration the discussion Butler provides of "reasonable self-love" in the first few chapters of his *Fifteen Sermons Preached at Rolls Chapel* (1726)[8], as well as the more systematic treatment of "rational egoism" given by Sidgwick in Book II of his *Methods of Ethics*. In my view, these two works provide the most complete analysis of prudence in British philosophy. They show that in the past there was no problem in justifying this line of behaviour as it was deemed perfectly rational. Moreover, Butler's and Sidgwick's analysis shows that Parfit's self-interest theory has the same structural features as the classical theory of prudence and therefore to reject S amounts to rejecting the classical theory. It is of vital importance to emphasize these similarities: only in this way is it possible to claim that Parfit's arguments refute the classical theory of prudence and therefore legitimize this author's historical importance.

2. *Butler's "reasonable self-love"*

The most detailed treatment by Butler of the topic of prudence is contained in his most important work on ethics, the *Fifteen Sermons Preached at Rolls Chapel*, first published in 1726, and republished in a second edition with an important new preface in 1729.

---

[8] Joseph Butler, *Fifteen Sermons Preached at the Rolls Chapel* (1726), with introduction, analyses, and notes by the Very Rev. W. R. Matthews, London, G. Bell & Dons LTD, 1967.

Butler's *Sermons* are presented as a treatment of interconnected topics and arguments mainly of ethical nature that have however often been developed unsystematically and are therefore hard to interpret.

Before making a direct examination of Butler's conception of "reasonable self-love" let us make a schematic overview of his conception of human nature, within which this doctrine is situated.

2.1. *The conception of human nature*

The importance of investigating the notion of human nature in the reconstruction of Butler's fundamental ethical theses is underlined by the author himself in the Preface to his *Sermons*:

> They were intended to explain what is meant by the nature of man, when it is said that virtue consists in following, and vice in deviating from it; and by explaining to show that the assertion is true.[9]

In the same Preface, Butler tells us that the principle according to which virtue lies in following nature is a very old one and has its origin in the ethical reflections of Stoic thinking. This principle, which the author believes still to be valid, requires a different proof from that which has been given in the past in which its psychological implications rather than its metaphysical implications are highlighted. In Sermons I, II, III and XI, Butler is engaged in the reconstruction of a conception of human nature on the basis of which virtuous behaviour is the only behaviour fully compliant with man's true constitution.

According to Butler, human nature consists of a plethora of internal principles that can easily be distinguished from each other, in spite of the philosophers' tendency to confuse them:

> Mankind has various instincts and principles of action, as brute creatures have; some leading most directly and immediately to the good of the community, and some most directly to private good.
>
> Man has several which brutes have not; particularly reflection or conscience, an approbation of some principles or actions, and disapprobation of others.[10]

[9] Joseph Butler, *Fifteen Sermons Preached at the Rolls Chapel*, Pref., p. 7.
[10] Joseph Butler, *Fifteen Sermons Preached at the Rolls Chapel*, Pref., p. 12.

The term "principle" is the most commonly noun used by Butler to refer indifferently to any internal source of human action. According to Butler's analysis, among the practical principles, it is important to make the distinction between "particular affections, appetites and passions" and "principle or general affection of self-love", and between the latter and the "natural principle" of benevolence. In addition to these elements it is possible to identify a principle of reflection which is used by mankind to express moral approval or disapproval of their actions, which Butler calls "conscience".

Self-love is identified with a general affection that urges us to act in conformity with our happiness; the general character of its object differentiates it from the other particular affections. It is also a rational principle, as it implies a capacity to distinguish between our present desires and our overall well-being. For the time being we shall not dwell on the analysis of this affection, to which the following section will be devoted.

The meaning Butler attributes to the principle of benevolence is more problematic. The status of this affection has caused a divergence among his commentators. The fundamental issue is whether this is a general principle, distinct from particular passions, or whether instead the term "benevolence" is simply a general term that refers indifferently to all the particular desires that have the welfare of others as their object[11]. It is beyond our present scope to clarify the different positions related to this problem. In agreement with T. Penelhum[12], I shall assume that the principle of benevolence must be interpreted as the "love of our neighbour". Benevolence is therefore a general desire whose object is not universal good but only the good of "that part of mankind, that part of our country, which comes under our immediate notice, acquaintance and influence, and with which we have to do".[13] Butler stresses that this general affection is also a rational principle: it actually demands the capacity to distinguish in others their short-term satisfactions from their long-term good.

The last principle identified in human nature by Butler is conscience. Unlike benevolence and self-love, conscience is not an affection. Butler presents it as "a principle of reflection in men, by which they distinguish be-

---

[11] For a detailed discussion of this problem see Terence Penelhum, *Butler*, London, Routledge & Kegan Paul, 1985. For a comparison of Butler's views on benevolence with those of Hutcheson and Hume see T. A. Roberts, *The Concept of Benevolence*, London, Macmillan, 1973; see also Amelie Rorty, *Butler on Benevolence and Conscience*, "Philosophy" 53 (1978), pp. 171-181. For a recent and extended study of Butler's ethics, see Stephen Darwall, *The British Moralists and the Internal 'Ought'*, Cambridge, Cambridge University Press, chap. 9.

[12] See Terence Penelhum, *Butler*, pp. 31-35.

[13] Joseph Butler, *Fifteen Sermons Preached at the Rolls Chapel*, Serm. XII, 3, pp. 186-187.

tween, approve and disapprove, their own actions".[14] The language used by Butler suggests that conscience is the only faculty by means of which human beings reflect on their own nature and on the practical principles and actions that comply with it. However, if what has been stated is correct, it is obvious that also self-love and benevolence demand a certain capacity for reflection as, in order to function correctly, each requires an awareness either of one's own nature and one's own needs or of that of the others and their needs. Therefore, while it is true that conscience must have a distinctive nature, this nature must reside in that special way in which it exerts its reflective power. In Sermon I, Butler strongly emphasizes that the distinctive nature of its judgments depends on the form they take on: namely, that of approval or disapproval. Conscience is therefore that faculty by means of which we approve or disapprove of our actions and the practical principle through which we act. In this sense, conscience can bridle self-love and benevolence when they lead us to perform actions that, in spite of our intentions, are contrary to our long-term interests and those of others.

The author of the *Sermons* believes that the empirical evidence in support of the presence of these principles in human nature is sufficient *per se* to demonstrate that human beings are not only liable to neglect the interests of others, but may also neglect their own. In this sense, promoting the good of others could be in agreement with the constitution of human nature to the same extent as acting to pursue one's own personal good.

However, even if this were sufficient to demonstrate that virtuous actions are to some extent compliant with human nature, it does not amount to claiming that virtue is compliant with our nature in a way that vice is not. To support this thesis it is necessary to attribute a well-defined meaning to the notion of "following nature". Butler lists three possibilities. It may be done by acting in compliance with one of the internal principles; or else simply by following the principle that, at some particular moment in time, has greater force than the others; lastly, we may follow our nature, acting in a way that is compliant with our entire constitution. Butler believes that only in the latter meaning it is possible to defend the thesis that virtue is compliant with human nature.

What does the concept of "entire constitution" of human nature refer to?

In Butler's conception, human nature, unlike that of the other animals, is organized in such a way that its principles have a superiority over particular inclinations which differs from that deriving from the simple intensity of the motivating force: their superiority is due to their status. Acting unnaturally simply means following a particular affection that one of the superior principles under those particular circumstances would ask us not to follow.

---

[14] Joseph Butler, *Fifteen Sermons Preached at the Rolls Chapel*, Serm. I, 8, p. 38.

Butler illustrates the point by referring to the unnaturalness of those actions we perform when we are in the sway of a violent desire, even though we know it will lead us to our doom:

Now what is it which renders such a rash action unnatural? Is it that he went against the principle of reasonable and cool self-love, considered *merely* as a part of nature? No: for if he had acted the contrary way, he would equally have gone against a principle or part of his nature, namely, passion or appetite. But to deny a present appetite, from foresight that the gratification of it would end in immediate ruin or extreme misery, is by no means an unnatural action: whereas to contradict or go against cool self-love for the sake of such gratification, is so in the instance before us. Such an action then being unnatural; and its being so not arising from a man's going against a principle or desire barely, nor in going against that principle or desire which happens for the present to be strongest … There must be some other difference or distinction to be made between these two principles, passion and cool self-love, than what I have yet taken notice of. And this difference, not being a difference in strength or degree, I call a difference in *nature* and in *kind*. And since, in the instance still before us, if passion prevails over self love, the consequent action is unnatural; but if self-love prevails over passion, the action is natural: it is manifest that self-love is in human nature a superior principle to passion. This may be contradicted without violating that nature; but the former cannot. So that, if we will act conformably to the economy of man's nature, reasonable self-love must govern. Thus … we may have a clear conception of the *superior nature* of one inward principle to another; and see that there really is this natural superiority, quite distinct from degrees of strength and prevalency.[15]

As Butler emphasizes in this passage, the question at stake in the case of the unnaturalness of serious rashness is whether to satisfy this impulse under those circumstances is contrary to the dictates of "self-love", a practical principle that, in accordance with the hierarchical structure of human nature, exerts its authority over any present inclination.

Butler claims that conscience has a supreme authority over all the other practical principles. In this sense, it may be claimed that acting against one's nature means following a lesser principle rather than the authority of conscience.

The doctrine of the natural authority of conscience is Butler's best known contribution to ethical theory, even though it is the one that caused his interpreters the greatest difficulty. In particular, it is not clear to what extent

[15] Joseph Butler, *Fifteen Sermons Preached at the Rolls Chapel*, Serm. II, 11, pp. 55-56.

the superiority of the conscience demands behaviour that may clash with the prescriptions of self love. In several passages, indeed, Butler seems to accept the thesis that moral virtue coincides with our own interests, if not actually to defend the much stronger thesis that to follow the dictates of conscience is justified solely by the fact that self love prescribes the same course of action.

2.2. *"Reasonable self love" and particular affections*

Before going on to analyse the specific nature of the self-love principle and the differences between it and particular affections, it is necessary to make a few further considerations regarding Butler's moral psychology.

Butler claims that to have an affection means having a particular goal, and that such a goal must be considered the object of that affection:

> Now, as reason tends to and rests in the discernment of truth, the object of it; so the very nature of affection consists in tending towards, and resting in, its objects as an end. … If we have no affections which rest in what are called their objects, then what is called affection, love, desire, hope, in human nature, is only an uneasiness in being at rest; an unquiet disposition to action, progress, pursuit, without end or meaning.[16]

The notion of the object of an affection thus finds its primary use in those cases in which the affection concerned has a particular aim or objective. The paradigmatic cases are presumably those in which the latter is a desire, of which appetites are special cases[17]. Other possible examples could be those of passions that are not desires, such as anger, resentment or compassion, in which however the psychological state under scrutiny is itself logically linked to the desire to do good or evil to someone.

The notion of the object of an affection allows us to appreciate the difference between particular impulses and self love. In his Preface to *Sermons*, Butler introduces the self-love principle, contrasting it with affections. The true contrast, however, is not with affections as such but with the particular nature of their objects. Self love is a general principle, not a special one; its general nature derives from its specific object, namely, happiness, or the long-term good of the subject. Butler writes:

---

[16] Joseph Butler, *Fifteen Sermons Preached at the Rolls Chapel*, Serm. XIII, 5, pp. 206-7.

[17] According to Butler's terminology, appetites are desires related to physical survival or well-being: such as hunger, thirst or sexual desire.

[…] private happiness or good is all which self-love can make us desire, or be concerned about: in having this consists its gratification; it is an affection to ourselves; a regard to our own interest, happiness and private good: and in the proportion a man hath this, he is interested, or a lover of himself […].[18]

Butler does not dwell at length on discussing what happiness represents for human beings; this is probably due to the general nature of self love, which excludes too precise an understanding of its object.

Happiness or satisfaction consists only in the enjoyment of those objects, which are by nature suited to our several particular appetites, passions and affections.[19]

Whatever the particular choices of each person that lead to happiness, a happy life will be one in which the majority of particular appetites and passions are satisfied in the long term.

The partial characterization of happiness provided by Butler helps clarify the status of self love. As several scholars have emphasized, the latter is a second order practical principle: that is, it is a desire to satisfy several desires and for certain objects of aversion to be removed in the long term.[20]

According to Butler, from this characterization it certainly emerges that self love is an affection; indeed it is a desire, although it may also be inferred that it is a rational principle. There are two reasons for this. First, self love demands the capacity to distinguish between a second order general object, such as the overall happiness of the individual, and the particular objects of passions. Second, since it may be exercised only through the judgment that particular objects will contribute to a certain extent to the general object of self love, it implies both the capacity to predict the effects of satisfying the various desires and the ability to calculate and compare the hedonic intensity associated with this satisfaction. Butler stresses the rational dimension of this principle by calling it "calm self love" and "reasonable self love". The rationality of prudent behaviour is highlighted also by the adjectives Butler uses to criticize those who are unable to achieve it: madness is often attributed precisely to those persons who are unable to master the complexity of their temporal existence, that is, who fail to compare several different satisfactions in separate times among themselves.

Butler also claims that it is precisely from the difficulty involved in achieving this line of behaviour that it follows that prudence is a virtue. In-

[18] Joseph Butler, *Fifteen Sermons Preached at the Rolls Chapel*, Serm. XI, 8, p. 169.
[19] Joseph Butler, *Fifteen Sermons Preached at the Rolls Chapel*, Serm. XI, 9, p. 170.
[20] T. Penelhum, *Butler*, chap. 1.

deed, insofar as it represents that reasonable self love, whose goal is our interest, prudence certainly does not coincide with immediate and particular passions. It requires the capacity to achieve proper behaviour in temporal matters: that is, to assess the consequences of one's actions beyond their more immediate effects, avoiding all actions based solely on the attainment of momentary satisfaction and, in Butler's words, not to fall into the error of forgoing a greater temporal good for a lesser one, that is, forgoing what is in our overall interest for the sake of momentary satisfaction.

2.3. *"Reasonable self love" and psychological egoism*

I want now to subject to further analysis the peculiar characteristics of the conception of prudence or the "reasonable self love" described by Butler, highlighting the differences between it and the doctrine of psychological egoism.

Psychological egoism is a thesis asserting that all actions are motivated by personal interest[21]. Two characteristics emerge from this highly schematic and general presentation. Psychological egoism is a thesis describing human nature which sets out to be empirically informative. Furthermore, this thesis is universal since it is aimed at characterizing the motives of all agents in all circumstances of action. It follows from the conjunction of these two characteristics that the conception in question might be refuted empirically.

Butler interprets the conception of psychological egoism using the language of his moral psychology: on the basis of his description, this conception asserts that all actions are performed under the influence of the affection of self love, that is, all motives can be reduced to the pursuit of one's own happiness. Such a theory has dangerous moral implications as it denies the existence of genuinely altruistic motivations. In his Sermon XI Butler will demonstrate how this doctrine must be rejected; it actually fails to stand up to the test of experience.

The first objection to psychological egoism asserts that this conception may be defended only on the basis of an erroneous interpretation of the genuine doctrine of self love. Acting in accordance with "reasonable self

[21] Psychological hedonism may be considered a special case of this thesis; it reduces all our motives to a desire for pleasure. In Sermon XI Butler devotes ample space to refuting this conception. It is due to two confused inferences: (1) the first infers from the fact that all my motives are *mine* the conclusion that the object of my desires must always be an internal state of mine, (ii) the second infers from the fact that I usually obtain pleasure from the satisfaction of my desires that my desires are always directed towards my own pleasure.

love", as already emphasized previously, means pursuing those lines of action the effects of which will presumably contribute to the overall happiness of the person. However, this in no way implies that the actions are due directly to the desire for one's own happiness. As mentioned earlier, the desire for happiness or personal well-being is a regulatory or second-order desire, the task of which is to ensure the best possible mix of satisfaction of one's particular passions. The normal function of this desire thus seems to be to approve or disapprove of particular desires in proportion to their capacity to contribute to the agent's overall happiness. It follows that this general desire cannot be a substitute for the function performed by the specific desires that are the direct cause of the actions. It seems quite apparent that, on the basis of this explanation of how self love can and should function, it is possible to come up with an immediate refutation of the doctrine of psychological egoism. Butler infers two different critiques from this.

First, it is possible to reject the thesis that our sole motivation is always the pursuit of happiness as a whole. Indeed on the basis of the explanation given above it follows that, in those circumstances in which we act in a fashion compliant with self love, there is an additional motive that the pursuit of one's own happiness encourages or at least allows to be satisfied. Butler points out how, in certain particular circumstances, it is actually possible that the pursuit of our own good is the only motive. On those occasions on which we refrain from doing something we actually desire because, for example, we deem it to be contrary to our interests, Butler asserts that self love is probably the only principle acting. However, these circumstances are rare and cannot be used as a suitable foundation for a universal theory of human motivation. Second, it is possible to reject the assumption, shared by many forms of psychological egoism, according to which the self love doctrine is incompatible with the existence of altruistic actions motivated by desires for the good of others. The essence of the discussion again hinges on the possibility of distinguishing self love and particular passions on the basis of their objects. As we saw, Butler claims that the objects specific to particular passions are particular desired objects or events, such as for example my lunch, a home, a holiday, etc. According to this distinction, he claims that although we might not be able to pursue happiness without first deciding whether our particular desires contribute to its attainment, we could act under the impulse of a particular desire without considering whether it contributes to our happiness. Once this corollary has been accepted there is no contradiction in imagining that the particular desires may include desires that contribute indirectly to the well-being of others, or actually have the good of others as their specific object. Whether these desires contribute to our happiness is a question that may be answered only through the very exercise of self love and to which no a priori answer may be given. But we can

thus conclude that desires for the good of others are not in principle more incompatible with self love than any other particular desire.

Butler's second argument is based on the observation that the substantive objective of self love, namely happiness, is attained more easily if we do not act under its impulse. He points out that in many circumstances we are better able to satisfy our happiness by not considering whether and to what extent what we wish to do contributes to it since the calculation of one's advantages can be an impediment precisely to those activities demanded by happiness:

Disengagement is absolutely necessary to enjoyment: and a person may have so steady and fixed an eye upon his own interest , whatever he places it in, as may hinder him from *attending* to many gratifications within his reach, which others have their minds *free* and *open* to.[22]

In this passage Butler points out that enjoyment is a form of attention and reflecting on self love may be a distraction. This is how he explains the fact that the persons who are always busy calculating their happiness are often unhappier than the others.

This refutation of psychological egoism has the merit of highlighting several important aspects of the theory of "reasonable self love". It actually shows not only that altruistic desires are not incompatible with the pursuit of one's own happiness but, more in general, sheds light on the risks linked to the adoption of a constant inclination to calculate the different satisfactions involved, which could lead to less satisfactory outcomes from the point of view of overall personal happiness itself.

3. *"Egoistic hedonism" in Henry Sidgwick's* Methods of Ethics

The idea that "calm self love" must be considered a normative principle, as it prescribes that individuals should pursue their overall happiness by prohibiting the satisfaction of those present inclinations which may prove detrimental in the long term, is picked up again by Henry Sidgwick in his *Methods of Ethics*. He includes "rational egoism"[23] among the methods of ethics, that is, among the rational procedures used by human beings to govern their behaviour whenever they seek to work out a complete synthesis of practical maxims.

---

[22] Joseph Butler, *Fifteen Sermons Preached at the Rolls Chapel*, Serm. XI, 9, p. 171.

[23] Following Sidgwick's use, the terms "egoistic hedonism" and "rational egoism" will be considered synonymous.

The inclusion of "rational egoism" among the methods of ethics will lead to the failure of the foundationalist project of ethics which represented one of the principal objectives of the *Methods*. In the well-known chapter devoted to philosophical intuitionism, in which the cognitive intuitionist epistemological framework forming the background to his treatment is outlined, Sidgwick shows that two mutually incompatible principles underpin prudence and benevolence, both of which are however self-evident: consequently the egoist could coherently maintain his own position without it being possible to refute it rationally.

Before presenting several of the characteristics of rational egoism, a few general considerations will be made concerning the philosophical project of the *Methods*.

3.1. *The objectives of the* Methods of Ethics

By the expression 'method of ethics' Sidgwick means any rational procedure by means of which it is possible to determine what human beings as single individuals must do. One of his strongest convictions is that common sense morality embodies different methods:

> Still I think that when a man seriously asks 'why he should do' anything, he commonly assumes in himself a determination to pursue whatever conduct may be shown by argument to be reasonable [...] And we are generally agreed that reasonable conduct in any case has to be determined on principles [...] But when we ask what these principles are, the diversity of answers which we find manifestly declared in the systems and fundamental formulae of professed moralists seems to be really present in the common practical reasoning of men generally [...].[24]

In Chapter 1 of Book 1, Sidgwick briefly discusses a variety of methods and principles which are linked in different ways and through different factual assumptions. More precisely, as J.B. Schneewind[25] emphasizes, Sidgwick, by analysing human moral reasoning, had identified two types of methods of ethics: 1) methods logically linked to the ultimate principles, and 2) methods indirectly linked to the ultimate principles. A method logically linked to an ultimate principle requires the moral agent to identify the

---

[24] Henry Sidgwick, *The Methods of Ethics*, p. 6.

[25] Jerome B. Schneewind, *Sidgwick Ethics and Victorian Moral Philosophy*, Oxford, Clarendon Press, 1977, cap. 6, pp. 194-98. See also, J. Schneewind, *Sidgwick and the Cambridge Moralists*, in Bart Schultz (ed.), *Essay on Henry Sidgwick*, pp. 93-121.

action to be performed exclusively through the only property that renders the actions right (right-making property). On the other hand, in a method indirectly linked to the ultimate principle, the moral agent identifies the actions to be performed not through the sole right-making property but by means of a characteristic linked to the latter through a contingent link (a criterial property). As Sidgwick himself asserts, his treatment was concerned solely with the "critical exposition of the different 'methods' … which are logically connected with the different ultimate reasons widely accepted".[26] The reason for this restriction is probably to be sought in the fact that Sidgwick was aware that one of the main causes of disagreement among human beings concerning their specific moral judgments consists of the differences related to their psychological, religious or metaphysical beliefs. By insisting on this restriction the moral philosopher was able to eliminate all the difficulties pertaining to realms of thinking that lay beyond the scope of ethics to investigate.

Among the many methods that are cloaked in varying degrees in the ambiguity of our moral language, Sidgwick claims that the following three methods can be distinguished: "egoistic hedonism", universalistic hedonism or utilitarianism, and intuitionism. He asserts the widely accepted commonsense view that it is rational to act both for one's private happiness as a whole and for the general happiness of all individuals. In this way it is easy to generate both the method of egoism and that of utilitarianism.

The intuitionist method, unlike the other two, is not linked directly to an ultimate principle. For the sake of simplicity intuitionism could be defined as the theory of ethics which considers as the ultimate aim of moral actions their compliance with certain unconditionally prescribed rules or dictates, without any consideration of the further consequences. The use of the term "dictates" implies including in this method the position according to which the ultimately valid moral imperatives are those referring to particular acts. Sidgwick himself, in Chapter 8 of Book I, alerts the reader to the different meanings he will assign to the term intuitionism, where those differences are due to the different generality of the intuitive beliefs recognized as ultimately valid.

The three methods analysed in *Methods* are not examined historically, as they are decision-making procedures that have effectively been proposed to govern everyday conduct, seeking to identify the changes that have come about over the centuries. Rather they are analysed insofar as, at least to the extent to which they are not mutually reconcilable, they represent alternatives from which human thought seems necessarily obliged to choose when

[26] Henry Sidgwick, *The Methods of Ethics*, p. 78.

it seeks to work out a complete synthesis of the practical maxims by striving to act in a perfectly consistent manner.

If, as it is often the case, the different common-sense methods applied in concrete circumstances provide mutually conflicting prescriptions, not all of them are acceptable:

[…] whereas the philosopher seeks unity of principle, and consistency of method at the risk of paradox, the unphilosophic man is apt to hold different principles at once, and to apply different methods in more or less confused combination […]. For if there are different views of the ultimate reasonableness of conduct, implicit in the thought of ordinary men, though not brought into clear relation to each other […] we cannot, of course, regard as valid reasonings that lead to conflicting conclusions; and I therefore assume as a fundamental postulate of Ethics, that so far as two methods conflict, one or other of them must be modified or rejected.[27]

Much of book IV of *Methods* is devoted to the attempt to harmonize and reduce to unity the different methods of ethics. However, I should like to point out that in the present article, in view of the objectives illustrated above, I shall not take into consideration the successful reconciliation between intuitionism and utilitarianism; instead, in view of the reconciliation between these two methods, I shall dwell on the problems raised by the attempt to seek a synthesis between utilitarianism and "rational egoism". Before directly addressing the problems linked to the relationship between these two methods, it is necessary to say something about the specific characteristics of "rational egoism".

3.2. *"Egoistic hedonism"*

Sidgwick devotes book two of *Methods* to the examination of "egoistic hedonism". He defines "egoism" as a method for determining the reasonable behaviour whereby each individual is supposed to adopt personal happiness as his own exclusive goal. Right from the outset, Sidgwick is aware of the innovative nature of the assumptions on which his investigation is based:

It may be doubted whether this ought to be included among received "methods of *Ethics*"; since there are strong grounds for holding that a sys-

[27] Henry Sidgwick, *The Methods of Ethics*, p. 6.

tem of morality, satisfactory to the moral consciousness of mankind in general, cannot be constructed on the basis of simple Egoism.[28]

He nevertheless deems it easy to dispose of this objection based on common-sense assertions that the principle has been widely accepted that it is reasonable for men to act in the way more likely to lead to their personal happiness:

Indeed, it is hardly going too far to say that common sense assumes that 'interested' actions, tending to promote the agent's happiness, are *prima facie* reasonable: and that the *onus probandi* lies with those who maintain that disinterested conduct, as such, is reasonable.[29]

According to the definition proposed by Sidgwick, it is necessary to define as egoistic the agent that, when faced with several possible lines of action, ascertains as accurately as possible the amount of pleasure and pain that is likely to result from each action and chooses the one which she believes will bring her the greatest happiness. The quantitative characterization of the rational goal of egoistic conduct deserves further clarification. The notion of the greatest possible happiness cannot be fully understood unless the meaning of "good on the whole" is clarified. A person's "good on the whole" is what she would desire and seek to achieve if she had fully understood all the consequences of all lines of conduct available to her. As Sidgwick perceptively points out, it is a terrible error to define a person's good simply as what would be desired if the outcomes of a given action could be predicted. It might always be possible that the choice of a particular object, while not emerging as an apparent good, that is, not different from what had been imagined, could on the whole be a bad choice owing to the concomitant aspects and long-term consequences. Sidgwick asserts that:

For it is not even sufficient to say that my Good on the whole is what I should actually desire and seek if all the consequences of seeking it could be foreknown and adequately realized by me in imagination at the time of making my choice. No doubt an equal regard for all the moments of our conscious experience – so far, at least, as the mere difference of their position in time is concerned – is an essential characteristic of rational conduct. But the mere fact, that a man does not afterwards feel for the consequences of an action aversion strong enough to cause him to regret it, cannot be accepted as a complete proof that he has acted for his 'good on the whole'. In-

[28] Henry Sidgwick, *The Methods of Ethics*, p.119.
[29] Henry Sidgwick, *The Methods of Ethics*, p. 120.

deed, we commonly reckon it among the worst consequences of some kinds of conduct that they alter men's tendencies to desire, and make them desire their lesser good more than their greater [...].[30]

Sidgwick claims that the principle prescribing that "one ought to aim at one's good on the whole"[31] must be considered as the self-evident intuition that underlies the "rational egoism" method. This principle is seen to be immediately self-evident when we consider individual goods of the person as similar parts of a quantitative or mathematical complex. In this perspective, the values of the individual goods will be assigned solely from the point of view of her maximum overall good, and the importance assigned to an individual good will be no greater than that which it has in the economy of her overall good. In other words, this principle states that a person must have an impartial interest for all parts of her conscious life. Of course, Sidgwick does not mean that a present good cannot reasonably be preferred to a future good on the strength of its greater certainty; he merely means to affirm that the mere difference of priority and posterity in time "is not a reasonable ground for having more regard to the consciousness of one moment that to that of another".[32]

Given Sidgwick's eudemonistic or hedonistic interpretation of the good, the principle of prudence may be expressed by stating that it is reasonable to forgo a present pleasure or present happiness in return for greater future pleasure or happiness or, more simply, that "a smaller present good is not to be preferred to a greater future good".[33]

From the foregoing the normative nature of the "egoistic hedonism" method emerges clearly: it consists in restricting a present desire in the wake of predictions of the more distant consequences deriving from such gratification.

The entire first chapter of book II is devoted to clarifying the notions of "interest" and "happiness", terms that in the author's opinion are too vague and ambiguous to be used in a scientific discussion on ethics. Sidgwick defines the notion of "greatest possible Happiness" as the "greatest attainable surplus of pleasure over pain"[34], where the term pleasure is used in its broader acceptance which includes all kinds of agreeable feelings: " the most refined and subtle intellectual and emotional gratifications, no less than the coarser and more definite sensual enjoyments"[35]. Acceptance of this quanti-

[30] Henry Sidgwick, *The Methods of Ethics*, p. 111.
[31] Henry Sidgwick, *The Methods of Ethics*, p. 381.
[32] Henry Sidgwick, *The Methods of Ethics*, p. 381.
[33] Henry Sidgwick, *The Methods of Ethics*, p. 381.
[34] Henry Sidgwick, *The Methods of Ethics*, p. 120.
[35] Henry Sidgwick, *The Methods of Ethics*, p. 127.

tative definition of the aim of egoism would imply that pleasures must be sought in proportion to their pleasantness, in such a way that the less pleasant state of consciousness cannot be preferred to the more pleasant state simply because the latter possesses some other qualities.

This conception of pleasure, which revisits Bentham's thesis of the complete homogeneity of pleasurable states of consciousness, completely contradicted the idea, defended by John Stuart Mill in his *Utilitarianism*[36], that it is possible to make a clear-cut distinction between qualitatively superior and qualitatively inferior pleasures. Sidgwick remarks:

This position, however, seems to many offensively paradoxical; and J. S. Mill in his development of Bentham's doctrine thought it desirable to abandon it and to take into account differences in quality among pleasures as well as differences in degree.[37]

According to Mill, differences in value between lower and higher pleasures were an "unquestionable fact"[38]. Sidgwick believed that the outlook defended by Mill could be accepted only if all the distinctions of quality could be resolved into considerations of quantity:

Now here we may observe, first, that it is quite consistent with the view quoted as Bentham's to describe some kinds of pleasure as inferior in quality to others, if by 'a pleasure' we mean (as is often meant) a whole state of consciousness which is only partly pleasurable; and still more if we take into view subsequent states. For many pleasures are not free from pain even while enjoyed; and many more have painful consequences. … and as the pain has to be set off as a drawback in valuing the pleasure, it is in accordance with strictly quantitative measurement of pleasure to call them inferior in kind.[39]

Sidgwick also believed that if non-hedonistic reasons for the preference were introduced into the egoistic calculation it would no longer be possible to consider egoism an autonomous method of ethics. Should it be admitted that the quality of the pleasures must be considered as something distinct from their quantity, and that it could even prevail over them, "egoistic hedonism" would no longer be clearly distinguishable from intuitionism.

---

[36] John S. Mill, *Utilitarianism* (1861), edited by Roger Crisp, Oxford, Oxford University Press, 1998, especially chap. 2.
[37] Henry Sidgwick, *The Methods of Ethics*, p.94.
[38] John S. Mill, *Utilitarianism*, p. 56.
[39] Henry Sidgwick, *The Methods of Ethics*, p. 94.

Before concluding this short reconstruction of the treatment of prudence as it appears in the *Methods*, I should like to examine what Sidgwick considered to be the difficulties implicit in the application of this method to cases of real conduct.

The fundamental assumption underpinning this method, which is implicit in the idea of considering a greater surplus of pleasure over pain as the ultimate aim of the conduct, is that all pleasures and all pains have a precise degree of positive or negative desirability which is knowable by the agents. Can it be assumed that in actual experience these degrees of desirability can be given with such precision? If this were false would it be a decisive objection to prudence?

Another assumption is that our pleasures can be increased and our pain decreased by means of forecasting and calculation. Nevertheless, it could be claimed that the practice of observation and hedonistic calculation inevitably tends to decrease our pleasures, at least the more important ones. It would thus seem problematic to try and attain our greatest happiness by attempting to pursue it scientifically.

Let us consider the latter objection first. Following Butler, Sidgwick affirms that it is possible to detect a difference between "extra-regarding" impulses and those whose object is our pleasure.[40] He also stresses that the greater part of our pleasure derives precisely from the satisfaction of those desires whose goals are different from pleasure itself.[41] In view of these premises it is easy to imagine what implicit danger lurks in the attempt to systematize conduct according to the principle of egoism: impulse towards our own pleasure could absorb the mind to such a degree as to become incompatible with the flow of those disinterested impulses towards particular objects, the existence of which is necessary in order to attain to a high degree that happiness toward which the principle of "egoistic hedonism" tends. This conclusion, which Sidgwick calls the "fundamental paradox of hedonism", must not be considered a decisive argument against this method:

I should not, however, infer from this that the pursuit of pleasure is necessarily self-defeating and futile; but merely that the principle of Egoistic Hedonism, when applied with a due knowledge of the laws of human nature, is practically self-limiting.[42]

---

[40] Henry Sidgwick, *The Methods of Ethics*, p. 44, see also p. 51.
[41] Henry Sidgwick, *The Methods of Ethics*, p. 44.
[42] Henry Sidgwick, *The Methods of Ethics*, p. 136.

In other words, according to Sidgwick, the only conclusion that can be drawn from the "paradox" is that the same method to achieve the end towards which egoism tends demands that to some extent we must place it outside our view and do not tend directly towards it. Once this danger has been clearly perceived it is no longer a cause of difficulty in the practical attainment of hedonism. As Sidgwick says:

For it is an experience only too common among men, in whatever pursuit they may be engaged, that they let the original object and goal of their efforts pass out of view, and come to regard the means to this end as ends in themselves: so that they at last even sacrifice the original end to the attainment of what is only secondarily and derivatively desirable. And if it be thus easy and common to forget the end in the means overmuch, there seems no reason why it should be difficult to do it to the extent that Rational Egoism prescribes [...].[43]

In Sidgwick's view, more serious objections may be raised concerning the possibility of performing precisely and reliably the methodical calculation of pleasure and pain required in order to adopt the method of egoism. In the first instance, if pleasure exists only insofar as it is felt, the fundamental assumption of egoism on the basis of which each pleasure has a quantitatively defined and measurable intensity must remain an a priori assumption that is not subject to any empirical verification. It is actually possible to assign a measure to a specific pleasure only when it is compared with other pleasurable sensations, but since this comparison can take place only in the imagination, it can only be hypothetically affirmed that, should it be possible for certain sensations to be felt simultaneously, it would be seen that one is more desirable than another in a definite proportion.

Second, even if it is taken for granted that each of our pleasures and pains can be measured precisely, the problem remains of whether we are in a position to know these quantities exactly. Indeed, even assuming we have an extraordinary predictive imagination, we would have to assume that during the measurement various different conditions were satisfied: 1) the mind would have to be in a perfectly neutral state in order to imagine all types of pleasure without bias for or against some specific sensation; 2) our capacity to enjoy certain specific pleasures must not change over time; 3) the assessment of the hedonic value of a past sensation must not be subject to error; 4) when we make use of the experience of others there must not be any difference between their sensitivity to the different types of pleasure and ours.

[43] Henry Sidgwick, *The Methods of Ethics*, p. 137.

3..3. *The dualism of practical reason*

Sidgwick believed that the numerous critiques that may be made to egoism do not make up a sufficiently strong argument to refute this method. Despite the difficulties involved, people are able to calculate their own pleasures accurately enough to satisfy the needs of their own lives. In fact, the strength of the normative reasons provided by egoism is never challenged by Sidgwick and it is precisely their universally binding nature that determines the failure of the foundational objectives of the *Methods*. In order to illustrate this point it must be borne in mind that Sidgwick, starting from the realization of the failure of "Mill's test" in favour of utilitarianism, comes to the conclusion that it is necessary to follow a method that is the opposite of the inductive one. One of the principal themes developed by Sidgwick in his *Methods of Ethics* is the demonstration that the grasp of self-evident first principles is essential for the rational foundation of utilitarianism. The construction of the utilitarian principle requires explicit recourse to two self-evident axioms. These are necessary to account for the universalistic dimension of which its specific nature is composed. The term "universalistic hedonism", which Sidgwick frequently uses as a perfect synonym of "utilitarianism", has the precise function of underlining this characteristic of universality.

The two axioms are those referring to the "principle of reciprocity" and to the relationship between the part and the whole. The former of the two states that "whatever action any of us judges to be right for himself, he implicitly judges to be right for all similar persons in similar circumstances"[44]. In other words, Sigdwick's idea is that unless there are significant differences among the agents or in the circumstances of actions, the same conduct is both morally valid and universally binding. The second axiom is represented by a universalization of the principle of the "egoistic hedonism" examined in the previous section, in which the relationship between the part and the whole is applied "from the point of view of the Universe"[45]. In short, as the egoist will consider her individual goods from the point of view of her maximum overall good, so the universalist hedonist will view her own good and that of others "from the point of view of the universe", from which the good of the single individual is important only insofar as it con-

[44] Henry Sidgwick, *The Methods of Ethics*, p. 379.
[45] Henry Sidgwick, *The Methods of Ethics*, p. 382.

tributes to the overall good produced in the universe. As Francesco Fagiani emphasized, in this view, "the overall goods of individuals appear as parts of a whole [...] to which they are subordinate and in which, all contributions being equal, the identity of the individual source from which the increase in the universal good comes is in no way significant".[46]

If, accepting Sidgwick's proposal, we identify the good with the non moral value of "pleasure" or "happiness", and if we accept the two self-evident axioms, utilitarianism is fully founded.

However, as Sidgwick himself points out quite "dramatically", the second of the two axioms is actually made up of two principles, the second of which may be rationally rejected even if the first is accepted. The first principle by itself provides the foundation of the "rational egoism" theory; only the acceptance of the second principle, that is, the consideration of one's own overall good as a part of the overall good of the universe, allows the egoistic dimension of ethics to be transcended by the universalistic one.

Sidgwick concludes his *Methods of Ethics* by acknowledging the fact that no rational argument exists that is capable of convincing those who have accepted egoism to accept the utilitarian prescription.[47]

Much of the contemporary discussion aimed at founding utilitarianism may be viewed as an attempt to come up with arguments that would allow, within the second axiom, to bridge the gap between the first and the second principle. Those who insist on the "separateness of persons" can only reject the second principle of the second axiom. Those who intend to develop Sidgwick's project further would have to propose a radical reappraisal of the notion of person by defending a conception of personal identity that is much less compact than the traditional one. However, once we decide to follow this path we will be forced to reconsider the categorical distinction between prudence and altruism. The philosophical reflections of Derek Parfit will be decisive for this new direction.

4. *Derek Parfit's "Self-interest Theory"*

Parfit describes his Self-interest or S Theory as a theory of individual rationality in which each individual is assigned the substantive objective of pursuing those outcomes that, given her set of desires, would allow her life

[46] Francesco Fagiani, *L'utilitarismo classico. Bentham, Mill, Sidgwick*, Napoli, Liguori, 1999, p. 53.

[47] In the final chapter of his *Methods*, Sidgwick affirms that the only possible way would be to postulate the existence of a utilitarian God who realizes the harmony between utilitarianism and prudence.

to unfold in the best possible way. In order to appreciate the peculiarity of this theory, it might be useful to imagine we could know all the desires of all persons – past, present and future. Moreover, each desire indicates both the person that has the desire and the time of her life in which it occurs (now, yesterday morning, in twenty years' time). In view of the enormous quantity of information involved, what would the rational course of action be? In other words, what desires should we take into greater account in deciding what to do?

Theories of rationality have suggested different answers to these questions. For instance, they may disagree as to whether it is rational to consider only our desires, or whether our future desires are to have the same weight as the present ones. The "Self-interest Theory" assigns significance only to the agent's desires and deems that those of the others can only indirectly influence the deliberative process that culminates in action. To use the technical jargon used in *Reasons and Persons*, this theory is agent-relative (it assigns to each individual a different substantive aim). Each of one's own desires directly provides the agent with a reason for acting and at any given moment the best rational action is dependent on the balancing of the relative weights of each of the reasons generated by those desires.

In the wake of Sidgwick's view of prudence, Parfit affirms that the force of these reasons is dependent exclusively on the intensity of the corresponding desires and thus the time at which they are perceived has no influence: future desires, according to S theory, must in themselves have exactly the same weight as we assign to our present desires. In Parfit's words, S is a temporally-neutral theory. Future events will be of less significance only if they are less likely to occur, but this does not mean that they are assigned less weight solely because, if they do take place, it will be later in time.

Parfit believes that it is possible to conceive of three equally plausible versions of this theory which differently interpret the meaning of best outcome. According to the "Hedonistic Theory", for each individual the best outcome is the one that ensures the greatest happiness. The various versions of this theory put forward different conceptions of happiness and of the ways of measuring it. In accordance with the "Desire-Fulfilment Theory", what is better for each individual is that which satisfies her desires throughout his life. On the basis of the "Objective List Theory", some things are good for us even if we do not desire them and bad for us even if we do not fear them. Different forms of this version exist according to what we consider to be good or bad.

These three theories coincide to a certain extent: they all agree in including happiness and pleasure among the things that enhance our lives and unhappiness and pain among those that worsen it. Without constraining him to choose among the three versions, this fact allows Parfit enormously to

simplify his treatment of "Self-interest Theory" by permitting him to discuss the "Hedonistic Theory" exclusively.

4.1. *How the "Self-interest Theory" may be self-defeating*

Parfit believes that numerous arguments may be constructed for the purpose of testing the plausibility of a moral theory or a theory of rationality. Among these, the simplest consists in demonstrating that a theory is self-defeating: this argument actually requires making no particular assumptions and in some cases is able to demonstrate that a theory fails on its own terms and must therefore be rejected.

Nevertheless, in the case of many theories, being self-defeating is not the same as demonstrating that those theories are unacceptable or must be rejected. In some cases this argument simply shows that a theory needs to be revised or extended, while in others it is unable even to demonstrate such a weak conclusion. In this section we will examine the outcome of this argument in the case of the "Self-interest Theory" according to Parfit's treatment. It will be highlighted how, although S is self-defeating, this in no way signifies a negative outcome for this theory.

In his article *Prudence, Morality and Prisoner's Dilemmas* [48] and at greater length in the first part of *Reasons and Persons*, Parfit identifies four ways in which a theory may be self-defeating: 1) a theory T is "indirectly self-defeating at the individual level" when it is true that, whenever someone attempts to achieve the objectives assigned to him by *T*, the latter are actually achieved less well on the whole; 2) a theory T is directly self-defeating at the individual level whenever it is certain that, if a person successfully follows *T* (that is, he succeeds in performing the act that, among those available to him, he more successfully achieves the objectives assigned to him by *T*), by this very fact he will act in such a way that the objectives assigned to him by *T* are achieved less well than if it had not followed *T* successfully; 3) a theory T is directly self-defeating at the collective level whenever it is certain that, if we all follow T successfully, for this very reason we will act in such a way that the objectives assigned to each one by T will be achieved less well than if none of us had successfully followed *T;* 4) a theory T is indirectly self-defeating at the collective level whenever it is true that, in the case that several persons follow the objectives proposed by T, those objectives are achieved less well.

[48] Derek Parfit, *Prudence, Morality and Prisoner's Dilemma*, "Proceedings of the British Academy" 65 (1979), pp. 539-64.

Since the "Self-interest Theory" is not a code of collective conduct, but a theory of individual rationality, the fact that it is indirectly or directly self-defeating at the collective level cannot be considered an objection to it. Parfit thinks it is easy to demonstrate that the "Self-interest Theory" is indirectly self-defeating at the individual level: for most people it is true that even if they never choose the line of action leading to a worse outcome, it would certainly be worse to be inclined to pursue one's own interest exclusively; it might be better to adopt another attitude.

It is worth emphasizing that the attitude responsible for the objection should not be interpreted as a set of self-interested motives always encouraging purely egoistic actions. Parfit, like Butler and Sidgwick before him, stresses that it is possible to pursue one's own personal interest by means of actions performed under the influence of altruistic motives or motives that are not directly self-interested:

Suppose that I love my family and friends. On all of the theories people affects what is in my interests. Much of my happiness comes from knowing about, and helping to cause, the happiness of those I love [...]. Suppose that I know that, if I help you, this will be best for me. I may help you because I love you, not because I want to do what will be best for me.[49]

Taking these explanations into account, Parfit believes that the best way to describe what it means for persons to have the attitude to pursue their personal interest is to affirm that, although often acting in pursuit of other more specific desires, they never do what they believe is worse for them. If this is true, these persons will explain themselves more clearly not by saying they have a disposition to pursue their own interest but by saying instead that they have the disposition never to go against it.

Let us now describe how, for an individual who adopts this disposition, S may be indirectly self-defeating. This would happen whenever a person, without ever going against her own personal interest, suffered a worse outcome than if she had adopted some other disposition. Even when persons succeed in never doing what is worse for them, the fact of never being willing to sacrifice their own happiness could be worse. Changing their disposition could prove more advantageous for them.

The following is one of Parfit's better known examples. Kate is a writer. Her greatest desire is for her book to be successful. Since the quality of her book is so important for her, she loves her work and her life appears to smile at her. If her desire to write the book was weaker, her work would be boring and her life, on the whole, would be negatively affected. Nevertheless, Kate,

[49] Derek Parfit, *Reasons and Persons*, pp. 5-6.

under the effect of her strongest desire is led to work so frantically and for such long hours that she ends up feeling exhausted and sometimes very depressed. As she is aware of this state of affairs, she is convinced that by working less frantically her book might be less successful but she would be happier, thus avoiding these periods of severe depression. If she accepts the "Self-interest Theory", thereby acquiring the disposition not to go against her own interest, Kate will come round to the idea that she should not overwork as by so doing she would do herself harm. This is an obvious case in which S would be self-defeating. Indeed Kate would always be able to avoid working at such a frantic rate only by tempering the intensity of her desire. This would represent an even worse outcome in terms of personal interest, since in this case work would be more boring for her and her life would be negatively affected. In Kate's case it is therefore obvious that never sacrificing one's own egoistic disposition can make things worse.

In this example the "Self-interest Theory" is self-defeating in its hedonistic version. If we were to accept the "Desire-Fulfilment Theory", we could reject Kate's idea that overwork is the cause of her problem: by working so hard, even though she wears herself out and occasionally suffers from depression, she manages to improve her book's quality. In this way, she ensures that her greatest desire is more fully satisfied. According to this "Self-interest Theory" this is a more satisfactory outcome for her.

For those who do not accept the hedonistic version of S, Parfit invents a different case. Let us imagine being lost in the desert and chancing to meet someone who can lead us back home in exchange for a certain sum of money. Let us imagine that we are unable to pay immediately, and that we promise to reward our rescuer as soon as we get home. Lastly, let us assume that we are transparent, that is, that we cannot lie without being caught. Since it would be worse for us to have to pay the agreed reward, if we know we are never willing to go against our own interest, we will never keep our promise to pay. Since we are transparent, also our would-be rescuer is also aware of this, and abandons us in the desert. For us it would have been better to be trustworthy, that is, to have the disposition to keep our promise even when to do so would make matters worse.

In the two cases described by Parfit, if an individual has the disposition of never going against her own interest, she makes the outcome worse. Parfit claims that this is true for most persons, for most of their lives. The question is – does this mean that the S theory is intrinsically false? Is this a sufficiently strong argument to reject the "Self-interest Theory"?

4.2. *The "Self-interest Theory" is not intrinsically false*

The objection in question would be fatal to the "Self-interest Theory" if it prescribed that persons should adopt the disposition never to go against their own interest. However, this would be an unacceptable thesis.

Parfit's argument is constructed on three theses underpinning the "Self-interest Theory". S claims that "for each person, there is one supremely rational ultimate aim: that his life go, for him, as well as possible" (thesis S1). When applied to acts, S claims both that each of us has most reason to do whatever would be best for himself (thesis S2) and that is irrational for anyone to do what he believes will be worse for himself (thesis S3). From the above three propositions a fourth thesis may be derived concerning the rationality of dispositions, that is, the set of motives that the "Self-interest Theory" prescribes that each agent should adopt. The fourth thesis claims that each agent should try to have or seek to maintain the best possible motives in terms of self-interest, that is this set of dispositions about which it may be affirmed that there is no other one that is better for her to have (thesis S4).

It is sometimes very difficult to know whether a set of motives may be causally possible, or whether it is one of the best in terms of S. Parfit nevertheless claims that there are also many cases in which a person knows that it would be better for her if her motives were to undergo some change: for such persons it may be true, as has emerged in the two preceding cases, that never to be willing to sacrifice one's self interest can lead to worse outcomes. Furthermore, in cases in which the person knows how to produce such changes, the thesis S3 implies that for these persons it would be irrational not to produce it, and that it would instead be rational to seek to have another disposition.

What these sets of motives actually are is partly a question of fact and the details of the response differ according to the different persons and the different circumstances of their lives: what we know in advance is only that it would be better for some persons if they were occasionally to go against their own interest and were willing to do what is worse for them. The limiting case is that in which for a person, under certain circumstances, it would be better to try and become completely irrational.[50]

Parfit claims that the "Self-interest Theory", although not intrinsically false, may nevertheless be refuted by means of an argument that challenges its very rationality. Before reconstructing this objection let us examine the conception of personal identity on which it is based.

---

[50] See the well-known case of *Schelling's Answer to Armed Robbery* in Derek Parfit, *Reasons and Persons*,pp. 12-13.

4.3. *Derek Parfit's personal identity theory*

In this section a brief outline is given of the central elements of the discussion of personal identity which makes up part three of *Reasons and Persons*. The essential arguments of Parfit's conception largely follow in the wake of the theses illustrated in numerous previous articles, and in particular those of his well-known article *Personal Identity*(1971).[51]

The two polemical objectives in that article, the thesis that personal identity is perfectly determined (the questions bearing on the identity of persons allow of only "yes or no" answers) and that according to which "what counts" when survival is at stake is personal identity itself, actually represent the central focus of the comprehensive discussion in *Reasons and Persons*.

Parfit claims that our view of the nature of persons and their continuing existence over time can be schematically presented as two theses: 1) persons are individual and ontologically non -reducible facts, whose continuing existence over time does not depend on (that is, it is not made up of) the existence of empirical, physical or psychological facts. From this it may be inferred as a corollary that the existence of the same person in two different times is a fact that is always perfectly determinable. 2) The continuing existence of these individual entities, that is, their numerical identity, is "what counts" when we are considering questions involving our survival. Numerical identity is the only thing that can justify the special interest we have in our existence and our future well-being.

Using a surprisingly large number of procedures, Parfit endeavours to demonstrate that what we are inclined to believe is not what we should believe because common sense has "a false view of the nature of personal identity"[52]. As an heir to that antisubstantialist tradition that had its first defender in Locke and in Hume its strenuous supporter, Parfit is defending a reductionistic or complex theory of personal identity which aims at reducing any discourse on the nature of persons to a description of the relations among classes of mental states that can be described "impersonally", thereby eliminating all forms of reference to subjectivity, to the point of view of the first person. He consequently puts forward a criterion of personal identity according to which our continuing existence consists in the recurrence of a relation of psychological connectedness and/or continuity among states of consciousness ("Relation R").

---

[51] Derek Parfit, *Personal Identity*, "Philosophical Review" 53 (1971), pp. 3-27.

[52] See Derek Parfit, *Lewis, Perry and the Matters*, in A.O. Rorty, *The Identities of Persons*, Berkeley, University of California Press, 1976, pp. 91-107. See also Derek Parfit, *Reasons and Persons*, chap. 10.

As it will be attempted to explain in the following sections, two highly innovative theses are implied in Parfit's conception. In the first place, the fact that the psychological connectedness is a relation that allows of variations in intensity means that it is also possible for cases in which our identity is indeterminate to occur. In the second place, if this thesis is accepted, it must be assumed that it is the relation of continuity and psychological connection between my present states and the future ones rather than personal identity *per se* what justifies the special interest in our future well-being.

Parfit's proposal has been interpreted as one of the most radical attempts ever made to eliminate the subject-person from the basic elements of the world. This contributed to making Parfit's reflections an essential point of reference both for those participating in the analytical debate who are interested in the general image of the person and for those involved in the discussion on the criteria of personal identity. It must be stated from the outset, however, that these two lines of reflection concerning Parfit, from our point of view, take on a significant, albeit limited, role. This is because our main interest lies not so much in the discussion of the nature of the person or in the way in which an answer to this question accounts for the thousand and one puzzles of personal identity, but rather in the consequences that Parfit's theory of personal identity has on the classical theory of prudence.

Parfit actually claims that close relations exist among the nature of persons and their identity over time and our reasons for acting. Once our shared opinions concerning personal identity have been changed we must consequently modify some of our beliefs concerning what we have most reason to do: we must reappraise out beliefs concerning rationality.

4.4. *Locke's legacy: the psychological criterion of personal identity*

The contemporary debate on personal identity is often characterized as referring to the principles which allow us to establish, for instance, that the person appearing before us is the same as the one we previously knew, where the principles sought must not be understood as mere pragmatic criteria (as, for instance, when the identity of a subject is established using his fingerprints), but refer to the justification of our identification procedures. As emphasized by Harold W. Noonan, it is possible in this connection to speak of the "logically necessary and sufficient conditions for which a person identified at a given moment is the same person as that identified at another"[53]. Nevertheless, it should be stressed that in this kind of investigation it is im-

[53] Harold Noonan, *Personal Identity*, London, Routledge,1989, p. 2.

portant to explicitly state the link between the search for such criteria and the more general, but no less exacting, question referring to the nature of the person. As Derek Parfit correctly points out we are confronted with two closely related issues: 1) What is a person's nature? 2) What is it that makes a person at two different moments one and the same person, or more precisely what is it that necessarily implies the continuing existence of each person over time?

Parfit claims that an answer to the second question is at least in part an answer to the first: the necessary characteristics of our continuing existence over time actually depend on our nature. In our examination of Parfit's position, for the sake of the explanation we shall not follow the order dictated by the logical priority of these questions but will deal with the nature of the persons after having answered the question of what is implied by their continuing existence over time.

Parfit defends a particularly sophisticated version of what in the contemporary debate is commonly defined as a psychological criterion. In very general terms, this conception states that personal identity implies the continuity of memory. This idea seems *prima facie* plausible because, it is claimed, it is precisely memory which makes most people aware of their own continuing existence.

The origins of this conception may be traced back to John Locke who, in Chap. XXVII of Book II of his *Essay concerning Human Understanding,* in several pithy pages, addresses what some believe may be considered the first comprehensive discussion concerning the criteria of personal identity which allow one to speak of a unitary subject that is continuous over time. For Locke the only fact that counts is the existence of direct memory connections, that is, memories of past experiences.[54] Parfit partly modified this Lockean conception. First of all, he considers that should no memory connections exist between two persons, let us say, between X today and Y twenty years ago, a continuity of memory may subsist just the same. This would be the case when a chain of linked memories exists between X and Y. This is a fairly frequent occurrence for the majority of people: every day they have memories of experiences they had the day before. It thus seems plausible to imagine that one of the conditions to be able to affirm that two persons at different times are the same person is that continuity of memory exists between them. Secondly, Parfit claims that the Lockean conception

[54] John Locke, *An Essay Concerning Human Understanding* (1689), edited with an introduction by Peter H. Nidditch, Oxford, Clarendon Press, 1975. For a detailed discussion on the influence of Locke's seminal ideas on the contemporary debate on the self, see Raymond Martin & John Barresi (ed.), *The Rise and Fall of Soul and Self: An Intellectual History of Personal Identity*, New York, Columbia University Press, 2006.

would however have to be corrected so as to take into account other psychological facts. As well as memories there are also other forms of direct psychological connection that necessarily have some weight in a personal identity criterion, such as desires, beliefs that are conserved over time, the connection linking an intention to the subsequent action in which it is implemented, salient features of a character, etc. Parfit terms all these kinds of direct psychological links "psychological connections".

Once Parfit modified the Lockean position along these lines, he places at the centre of his psychological criterion the relation of psychological continuity. Parfit defines this fundamental relation as the occurrence of chains of strong psychological connections. Here *strong* is meant to signify the existence of connections that are acceptable on average; for instance, an adult person might be called upon each day to recall at least 50% of the previous day's experiences, and so on for all the other types of psychological connections mentioned above.

Psychological continuity, unlike psychological connection, is a transitive relation and may therefore represent the personal identity criterion over time. This enables us to formulate the "psychological criterion" of Parfit's personal identity: (1) "psychological continuity" exists only when there are linked chains of strong connections. X today is the same person as Y in a previous moment only if (2) X is in psychological continuity with Y. (3) personal identity over time consists precisely in the occurrence of facts like (2).

4.5. *The "Reductionist" conception of personal identity*

According to Parfit's psychological criterion, personal identity over time merely implies different types of psychological continuity. Parfit affirms that this conception may be considered as a "Reductionist" theory of personal identity. In the latter the fact of the identity of a person over time, that is, her continuing existence, is deemed to consist solely in the occurrence of simpler, i.e. psychological, facts. These facts may be described in an impersonal way, that is without explicitly affirming that these were the experiences of a specific person. In other words, it is possible to describe all the psychological facts characterizing the mental life of a person in purely objective terms, in the third person, thereby eliminating all references to a first-person point of view.

Opposed to this theory are the "Non-Reductionist" conceptions of personal identity. In their stronger version they affirm that personal identity over time does not consist solely in physical and/or psychological continuity: it is an additional fact, distinct from the latter. It consists in the existence of a spiritual substance, a simple purely mental entity that accounts

for both the unity of consciousness in the various moments in time and the unity of a life as a whole.

In the "Reductionist" conception, persons are merely sets of experiences made up of relations of direct "psychological continuity" or by weaker forms of connection. In accordance with Parfit's well-known metaphor, they resemble clubs: entities that exist in a certain sense, but which are not included among the substantial elements of the world, as entities characterized by being centres of experience, but which are completely exhausted in the individuals that constitute them.

If we are seeking an example of the ontological depotentiation of the subject, suffice it to examine Parfit's description of dying, which seems to resemble more closely the break-up of a meeting than the irreparable loss of something. For Parfit: "Instead of saying, 'I shall be dead', I should say, 'There will be no future experiences that will be related, in certain ways, to these present experiences'"[55].

4.6. *"Reductionist" thesis: "what matters" is not personal identity*

Within the framework of neurobiological and neuropsychological research, the results of the clinical examinations performed on patients suffering from different types of disorder seem to cast doubts on the conventional image of ourselves as unitary and continuous entities. The conflict between the philosophical considerations triggered by several clinical cases and the common sense intuitions concerning the self is a topic that receives extensive treatment in part three of *Reasons and Persons*. According to Parfit's interpretation, the forms of "dissociation of consciousness" believed to take place in the case of so called "split brains", those in which the connections between the cerebral hemispheres have been surgically severed, provide strong arguments in favour of his reductionist conception. They are deemed to demonstrate that "what matters", that is, what justifies the special interest we feel in our future, is not personal identity but the relation of psychological continuity and/or connection (relation R).

Parfit imagines a radical case of ramification of the streams of consciousness in which the brain of an individual A is split and transplanted into that of his two brothers B and C. The latter will have a relation of complete psychological continuity with their donor. Both of them, after waking up after the operation, will believe they are the dead brother: they will have the impression of remembering having lived his life, will have his same desires and his same intentions.

---

[55] Derek Parfit, *Reasons and Persons*, p. 281.

In this imaginary case, "Relation R" ("psychological connectedness" and/or "psychological continuity") is configured as a bifurcation. However, personal identity cannot take on this form. The donor and his two brothers, thus constituted, cannot be the same person. Since the donor cannot be the same as two different persons, and since it would be arbitrary to say that only one of the two brothers is the same as the first one, the best way of describing this case is to say that neither person is A.

Unlike ordinary cases, in which personal identity is merely the occurrence of "Relation R" (indeed in practically all real cases R takes on the form of a one-to-one relation: that is, it exists between a person who currently exists and a future person), these are cases in which "psychological continuity" and "psychological connectedness" exist without identity.

The question might thus be asked of whether the lack of identity is really so important. Parfit's answer is negative: what really counts is the "Relation R" whatever its cause (in normal cases the persistence of the "Relation R" is guaranteed by the continuity of the central nervous system, which is indeed the natural cause).

Parfit illustrates this point by discussing an imaginary story. Let us imagine that a 'Star Trek' science-fiction-like device is used to scan my body and break it up into its component parts and then sends a signal to Mars by means of which a body identical to the original is recomposed. Subjectively speaking what happens is this: I press a button on Earth and immediately find myself on Mars. Assuming total psychological continuity I could say that the individual recomposed on Mars is identical to my self on Earth. Let us imagine, however, that a second copy of myself is sent to Saturn. For the reasons given above it is no longer possible to assert that I am identical to the individual sent to Mars. On the other hand, it is easy to believe that this lack of identity does not count for very much: after all, the situation is the same as before with the addition of a third person on Saturn. A few problems could conceivably arise because of the split (quarrels over possessions, love for the same wife, etc.) although, according to Parfit, the type of survival that I am guaranteed by psychological continuity in the ramified case is what I ought to assign value to.

If we accept Parfit's "psychological criterion", each ramification corresponds to the death of an individual A and the birth of two "Parfitian heirs", B and C, in his place, neither of whom is identical to A. But as Di Francesco rightly points out, this is a death in which no one actually dies: "the subject is actually not an added value vis-à-vis the continuity of the

experience and if the latter persists (albeit is multiplied), we have no reason to complain of the loss of anything real"[56].

4.7. *Refutation of the classical theory of prudence*

Now I shall present a comprehensive treatment of the argument by means of which Parfit believes it is possible to refute the "Self-interest Theory".

In S the "requirement of equal concern" is fundamental: a rational person ought to have equal concern for all parts of his own future. This means that each of us may attribute less importance to what may happen in the future only if this remoteness makes the event less probable. According to Parfit this thesis may be challenged on the basis of the reductionist conception. As emphasized in the preceding section, on the basis of Parfit's position, what fundamentally matters is psychological continuity and/or connectedness. In more than one point Parfit reiterates that both these relations play an important role in determining the special interest we attribute to our future. With these premises in mind, Parfit states a general thesis:

(C) My concern for my future may correspond to the degree of connectedness between me now and myself in the future. Connectedness is one of the two relations that give me reasons to be specially concerned about my own future. It can be rational to care less, when one of the grounds for caring will hold to a lesser degree. Since connectedness is nearly always weaker over longer periods, I can rationally care less about my further future.[57]

It should be noted that Parfit defends a discount rate referring not to time but to the attenuation of one of the two relations making up what has fundamental importance. Unlike the discount rate referring to time, this new discount rate is unlikely to be valid for the near future.

According to Parfit we must accept the thesis (C). Even if there are some exceptions, numerous relations must be judged less important when they occur with reduced levels of intensity: friendship, complicity, kinship, responsibility are but a few of the possible examples. Psychological connectedness must be considered in a like fashion. If we accept (C) we are rejecting the requirement of equal concern. This requirement is central to the "Self-interest Theory" and so we must reject this theory.

---

[56] Michele Di Francesco, *L'io e i suoi sé. Identità personale e scienza della mente*, Milano, Raffaello Cortina Editore, 1998, p. 195.

[57] Derek Parfit, *Reasons and Persons*, p. 313.

Parfit imagines that a defender of the theory S could retort that it is possible to modify the theory in question in such a way as to take his objection into account by incorporating a discount rate referring to psychological connectedness. According to this revised version, the dominant interest of a rational being ought to be that referring to her own future, although at that moment she might have less interest in those parts of her future with which she currently has a less close connection.

Parfit counters this response by arguing that this revised theory could not be considered a version of the "Self-interest Theory". Indeed the revised theory severs the fundamental link between S and the person's good on the whole. In the previous sections it has been shown how a central characteristic of the various formulations of the classical theory of prudence is that it is irrational for anyone not to do what she believes to be her good on the whole. In the revised theory, on the other hand, this thesis would have to be abandoned: if it is not irrational to be less concerned with certain parts of one's own future, it may not be irrational to do what is deemed worse in relation to one's own good on the whole.

As can be seen from the latter statement, the reply by the defender of S cannot be accepted and Parfit's objection is still decisive.

4.8. *The immorality of imprudence*

The outcome of Parfit's argument against the "Self-interest Theory" shows that it is no longer possible, as required by classical theory, to consider imprudent actions as irrational, since prudence cannot be equated with practical rationality. This means not having any more philosophical arguments to criticize imprudent actions. As Parfit affirms:

> If we believe that an imprudent act is not irrational, the charge 'imprudent' will cease, for many people, to be a criticism. It will become merely descriptive, in the way that, for many, 'unchaste' is merely a description.[58]

It therefore becomes necessary to seek a new theory that, by using a criterion other than rationality, will enable us to censure imprudent actions. Parfit suggests modifying our moral theory in such a way as to extend its application also to those actions that, in the past, were not the primary object of moral evaluation.

[58] Derek Parfit, *Reasons and Persons*, p. 318.

In this perspective, Parfit believes two different strategies may be pursued. He limits himself to describing them in very general terms, without actually choosing between them.

The first proposal consists of an appeal to consequentialism, and in particular to an agent-neutral principle of beneficence. If, in order to obtain lesser benefits in the present, an individual acts in such a way as to obtain greater hardship in her old age, she acts in a way that, when considered impartially, is the cause of worse consequences as it increases the quantity of suffering in the world, in accordance with this line of argument it may thus be affirmed that this individual acts in a morally deplorable way as her imprudence makes the outcome worse. As Parfit claims in *Reasons an Persons*, from the impartial perspective of consequentialism, "it is no excuse that the outcome will be worse only for me"[59].

Conversely, the second strategy consists in extending that part of moral theory that is "agent- relative". This involves our special obligations towards those with whom we have special relations: those with children, parents, patients, clients, are only few examples. It could be affirmed, Parfit goes on to say, that the relation between me in the present and me in the future sets up similar special obligations.

Parfit is aware that a revision of our moral conception in one of these two ways "would be, for many people, a large change in their conception of morality", since it seems to be a very deep common belief in our shared ethical thinking that "it cannot be a moral matter how one affects one's one future"[60]. However, if we accept a complex and reductionist account of the self this is the only strategy available. From this perspective we can no longer maintain that there is a categorical distinction between the way our actions affect our future self and the way they affect other selves. The only reasons that apply to these situations are the moral ones.

[59] Derek Parfit, *Reasons and Persons*, p. 319.
[60] Derek Parfit, *Reasons and Persons*, p. 319.