Université de Montréal

The Authority of Moral Conscience in Joseph Butler's Ethics

par
Esmé A. Vlahos

Département de philosophie
Faculté des arts et des sciences

Mémoire présenté à la Faculté des études supérieures
en vue de l'obtention du grade de Maîtrise
en Philosophie

Décembre 2004

Université de Montréal
Faculté des études supérieures

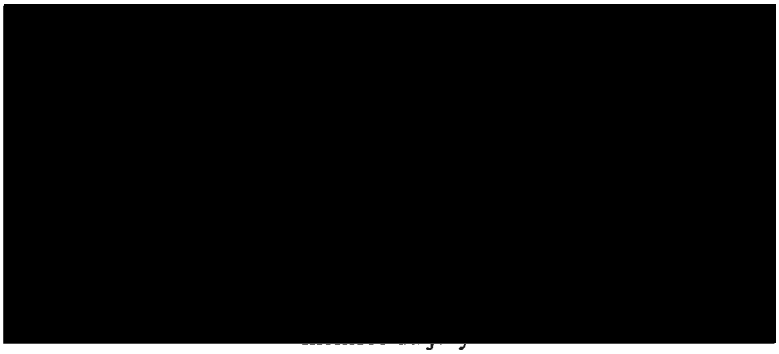
Ce mémoire intitulé :

The Authority of Moral Conscience in Joseph Butler's Ethics

présenté par :

Esmé A. Vlahos

a été évalué par un jury composé des personnes suivantes :



Résumé

Ce travail se propose d'étudier l'autorité de la conscience morale dans l'éthique de Joseph Butler. Dans le premier chapitre, nous examinerons les rôles des passions (en particulier de la compassion et du ressentiment), des principes rationnels de l'amour de soi et de la bienveillance ainsi que de la conscience, dans la psychologie morale de Butler. Dans le chapitre 2, nous présenterons sa réfutation de l'égoïsme psychologique et le « cool hour passage » des Sermons dans lequel il semble préconiser l'égoïsme éthique. Nous verrons que, selon Butler, la motivation humaine est variée et qu'il n'y a pas de contradiction fondamentale entre l'amour de soi et la bienveillance. Dans le chapitre 3, nous illustrerons la théorie de la conscience morale et son autorité suprême dans la hiérarchie des principes de la nature humaine chez Butler. Selon lui, notre nature est adaptée à la vertu, et la vertu consiste à suivre la nature. Finalement, dans le chapitre 4, nous exposerons les arguments en faveur de l'autorité de la conscience : (i) l'argument téléologique-fonctionnel selon lequel la finalité humaine est la vertu ; (ii) l'argument constitutionnel selon lequel la nature humaine est adaptée à la vertu parce que nous possédons une faculté morale suprême ; et (iii) l'argument autonomiste voulant que la conscience soit la condition même des raisons prépondérantes d'agir d'un agent moral. En conclusion, nous explorerons la possibilité que l'éthique de Butler puisse être séparée de sa théologie et nous nous demanderons comment ceci influe sur l'autorité de la conscience morale.

Mots clés

Philosophie, théologie, égoïsme, altruisme, amour de soi, bienveillance, vertu, nature, passions, compassion

Abstract

This essay proposes to study the authority of moral conscience in Joseph Butler's ethics. In Chapter 1, we will examine the roles of the passions (in particular compassion and resentment), of the rational principles of self-love and benevolence, and of conscience in Butler's moral psychology. In Chapter 2, we will present Butler's refutation of psychological egoism and the "cool hour passage" in his Sermons where he appears to advocate ethical egoism. We will see that, for Butler, human motivation is varied and there is no special contradiction between self-love and benevolence. In Chapter 3, we will illustrate Butler's theory of moral conscience, demonstrating its supreme authority in the hierarchy of principles of human nature. According to Butler, our nature is adapted to virtue and virtue consists in following nature. Finally, in Chapter 4 we will set forth the arguments in favour of the authority of conscience: (i) the teleological/functional argument that we are designed for virtue; (ii) the constitutional argument that our nature is adapted to virtue since we possess a supreme moral faculty; and (iii) the autonomist argument that conscience is the very condition of a moral agent's having overriding reasons to act. To conclude, we will explore whether Butler's ethics can be separated from his theology and how this affects the authority of moral conscience.

Key Words

Philosophy, Theology, Egoism, Altruism, Self-love, Benevolence, Virtue, Nature, Passions, Compassion

TABLE OF CONTENTS

Acknowledgments

I would like to express my gratitude to Luce Des Marais, Giselle Nyiri, Christine Tappolet and Margaret Ruth Watson for their helpful suggestions, support and encouragement.

Introduction

Bishop Joseph Butler (1692-1752) was an Anglican priest who first published in 1726 *Fifteen Sermons* on human nature he delivered during his eight years of service as preacher at the Rolls Chapel in London, England (hereinafter referred to as the *Sermons*[1]). As Alan Millar notes, "the *Fifteen Sermons* are the work of a moralist with the very definite aim of encouraging people to the practice of virtue." (in Cunliffe 1992, p. 293). Butler was intent on bringing to the forefront of the minds and hearts of the members of his congregation his own preoccupations with truth and "the important question, What is the rule of life?" (*Sermons*, The Preface, (1), p. 4). He challenged popular ways of thinking at the time, which were largely sceptical as to matters of morals, and attempted to reinstate belief in the virtuous life.

As far as Butler's stylistic aim is concerned, in the Preface to the *Sermons*, he writes that the study of morals requires a peculiar kind of attention in order to clearly articulate for the layman what is at stake. Contrary to works of entertainment and leisure, such "scientific" subjects as ethics require that we state things as we find them, sometimes in great detail. According to Butler, this involves separating complex issues which should not be confused. However, he was equally aware that his desire for utmost clarity could at times make questions appear obscure. Thus, he writes in the Preface to the *Sermons*: "However, upon the whole, as the title of *Sermons* gives some right to expect what is plain and of easy comprehension, and as the best auditories are mixed, I shall not set about to justify the propriety of preaching, or under that title publishing, Discourses so abstruse as some of these are; [...]"((10), p. 6).

In publishing the *Sermons*, Butler did not intend to present his reader with a complete, consistent theoretical ethical treatise. In his own words:

---

[1] References are to The Very Rev. W. R. Matthews 1967 edition's page numbers, as well as to its paragraph numbers (indicated in brackets) which follow John Henry Bernard's complete edition of Butler's works.

> It may be proper just to advertise the reader, that he is not to look for any particular reason for the choice of the greatest part of these Discourses; their being taken from amongst many others, preached in the same place, through a course of eight years, being in great measure accidental. Neither is he to expect to find any other connection between them, than that uniformity of thought and design, which will always be found in the writings of the same person, when he writes with simplicity and in earnest. (*Sermons*, The Preface, (45), p. 27).

Moreover, as a preacher, Butler was especially interested in what he called the "practical" aspect of human nature, *i.e.* what we would today call, in part, moral psychology, or the discipline that looks into questions bordering on psychology and ethics. The issue addressed by Butler is that of "the constitution of human nature", in other words, the answer to the question of how are we constituted or what is our true nature. Butler's query is normative as well as descriptive, for he not only explores the various components in human nature and how they interact, but further inquires into our purpose in life, given such a constitution. In so doing, he attempts to refute the theory that we are solely preoccupied with our personal survival and furthering our own interest, to the exclusion of the welfare of others. As we will see in our second chapter, in Butler's view such a way of thinking, now named *psychological egoism*, ignores the fact that there are other motivational sources in humans which are paramount and arises from confusion as to the meaning of the words "interest" or "interested". Butler is considered by many a writer in ethics as having succeeded in his attempt to refute psychological egoism and hedonism. We will examine his arguments against these theories as well as their normative version, *i.e.* the pursuit of our interest (or pleasure) as a morally recommended way of life or *egoism as a virtue*, sometimes called *rational* or *ethical egoism*[2].

---

[2] See Ayn Rand 1964; Robert Shaver 1999; and Charlie Dunbar Broad 1953 (in Cheney 1971) and Jan Österberg 1988, respectively.

The study of human nature and the refutation of psychological egoism bring Butler to the normative question of how we should conduct ourselves given such a nature. Butler is interested in the purpose human beings were designed to accomplish during their life, what we must do to properly fulfil our "natural" destiny. The language of ethics impregnates his discourse alongside natural theology, as Butler endeavours to reply to fellow Christians who ask themselves: "How should I live my life?", or its equivalent: "What is my duty?". Butler's moral philosophy incorporates a naturalistic approach, and in this regard, he does not stand apart from the most important 17th and 18th century British moralists. As Gerard J. Hughes observes:

> The classical moral philosophers in the Western tradition have almost all adopted some form of natural law theory. However, it must also be said that while they share the crucial tenet of a natural law theory, *i.e.* that moral duties can be ascertained by reflection on human nature, they differ widely in their views about what human nature is and, as a result, about the moral theory that can be derived from it. Thus, for example, Hobbes believed that human beings are motivated entirely by their desire for pleasure and aversion to pain, and that this entails that moral theory should be egoistic in character. Butler and Hume take a different view of human desires and hence reject a Hobbesian egoism. (G. J. Hughes 1986, pp. 412-413)

In his effort to ascertain our duty, Butler's point of departure is our daily existence and the common denominators observed in human nature. His thesis is inspired by "the ancient moralists" who claimed that we are "born to virtue" and that our nature is so constituted that virtue is our proper end. Butler thus sets forth the aim of the *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. That the ancient moralists had some inward feeling or other,

> which they chose to express in this manner, that man is born to virtue, that it consists in following nature, and that vice is more contrary to this nature than tortures or death, their works in our hands are instances. (*Sermons*, The Preface, (13) pp. 7-8).

The underlying tenet is that humans act naturally when they pursue virtue and that vice runs counter to our nature. However, Butler felt it necessary to explain this inner conviction of the ancients that our natural end is virtue, since in the 18th century it was no longer considered a certitude.[3] He states that his goal is to better understand the "system" of human nature, with all the various relations between the parts that constitute it, following in the steps of those before him who had written "treatises upon the passions" (*Sermons*, The Preface, (13), p. 8).

With this objective in mind, Butler decides that he will not use what he calls the "abstract" method of reasoning employed by other moral theorists, such as his teacher and correspondent Samuel Clarke. Rather, he chooses a method that starts by observing what is found in the world, and then attempts to clarify the role of constituents parts and construct from the relations between the parts their *raison d'être*. In ethics, this involves examining human nature closely to determine what should guide our conduct in order for us to be in harmony with our entire being and the end for which we were created. Butler describes this method as follows:

> There are two ways in which the subject of morals may be treated. One begins from inquiring into the abstract relations of things: the other from a matter of fact, namely what the particular nature of man is, its several parts, their economy or constitution; from whence it proceeds to determine what course of life it is, which is correspondent to this whole nature. In the former method the conclusion is expressed thus, that vice is

---

[3] *Sermons*, The Preface, (13), p. 9 with respect to William Wollaston's *Religion of Nature Delineated*: "A late author of great and deserved reputation says, that to place virtue in following nature, is at best a loose way of talk."

> contrary to the nature and reason of things: in the latter, that it is a violation or breaking in upon our own nature. Thus they both lead us to the same thing, our obligations to the practice of virtue; and thus they exceedingly strengthen and enforce each other. [...]
>
> The following Discourses proceed chiefly in this latter method. The three first wholly. (*Sermons*, The Preface, (12-13), pp. 6-7)

Though Butler does not exclude the validity of the "abstract" method of reasoning in ethics, he favours the approach which begins with a thorough exploration of our nature rather than a theoretical deduction of what is the true "nature and reason of things".

Methodologically, Butler also makes extensive use of arguments "by analogy". For example, with the help of the "analogy of nature", Butler hopes to reveal the merits of the virtuous life, without the necessity of founding them in Christian revelation. In 1736, Butler published *The Analogy of Religion, Natural and Revealed, to the Constitution and Course of Nature. To which are added, Two Brief Dissertations: I. On Personal Identity.–II. On the Nature of Virtue.* (hereinafter referred to as the *Analogy*, while the second dissertation will be referred to as the *Dissertation*)[4]. In *Part I. Of Natural Religion* of the *Analogy*, Butler induces from the course of nature, as a "scheme imperfectly comprehended" by us due to our ignorance of its laws, probable (as opposed to demonstrative) evidence for the following. From the existence of personal identity, Butler reasons to the probability of immortality of the soul and of an afterlife. From the natural government of the world by rewards and punishments, he induces the moral government of God and a future state of rewards and punishments. From our being to a large degree responsible for ensuring our own interest and happiness in the present life, he infers that, as

---

[4] References will be made to the Matthews 1967 edition for the *Dissertation* and to the Samuel Halifax 1838 edition for the *Analogy*.

immortal souls, we are in a state of trial or probation for a future life and are hence free agents capable of moral improvement by discipline and accountable for our acts. Once again, Butler proposes the practice of virtue in answer to the inquiry "of real, and of the utmost importance to us to have answered: the inquiry, What is our business here? The known end then, why we are placed in a state of so much affliction, hazard, and difficulty, is, our improvement in virtue and piety, as the requisite qualification for a future state of security and happiness." (*Analogy*, p. 113).

In the *Sermons* as well, Butler uses argument by analogy when comparing the "system" of human nature to that of a watch: just as a watch is a unit, comprised of specific relations between its parts and serving the end for which it was designed, *i.e.* to tell time, the moral self is a system, constitution, or economy of constituent parts. According to Butler, any treatise in ethics which neglects or exaggerates the importance of one of these parts, or of one of the relations between the parts, would be incomplete and therefore biased (like a badly written user's manual). Moreover, if we were to modify the relations between the components of a watch, or neglect one of its parts, it could not fulfil its true function. Likewise, the proper functioning of a human being reveals that the constituents of human nature stand in a relation to one another which permits us to pursue virtuous living. Butler's argument is clearly teleological: when we fail to pursue virtue or act contrary thereto, our moral "system" is "out of order", like a vending machine which retains the coins without distributing the product requested. Following Alan Millar and Stephen Darwall, we will call this the "teleological/functional" line of argument[5]:

> Every work both of nature and of art is a system: and as every particular thing, both natural and artificial, is for some use or purpose out of and beyond itself, one

[5] See Darwall 1995, p. 262 *ss.* and Millar, in Cunliffe 1992 at p. 294 and 1992, p. 488.

> may add, to what has been already brought into the idea of a system, its conduciveness to this one or more ends. [...] And from the idea itself it will as fully appear, that this our nature, *i.e.* constitution, is adapted to virtue, as from the idea of a watch it appears, that its nature, *i.e.* constitution or system, is adapted to measure time. (*Sermons*, The Preface, (14) pp. 9-10)

According to Butler, our nature is adapted to virtue due to our possessing an "authoritative" principle of reflection, *i.e.* moral conscience, which rules our constitution. We will see that Butler introduces at this point the notion of a hierarchy in the principles of human nature, with moral conscience at the top, self-love and benevolence in the middle, and the passions below. It is generally agreed that Butler's system is a triple-tiered pyramid, with the rational principles governing the passions and the moral faculty ruling all. Motivation for our actions is found on all three levels, but is "stronger" at the bottom than at the top. However, conscience is the only principle in human nature which possesses supreme authority over all others. Once again, following Millar and Darwall, we will call this the "constitutional" line of argument[6]:

> Appetites, passions, affections, and the principle of reflection, considered merely as the several parts of our inward nature, do not at all give us an idea of the system or constitution of this nature; because the constitution is formed by somewhat not yet taken into consideration, namely, by the relations which these several parts have to each other; the chief of which is the authority of reflection or conscience. (*Sermons*, The Preface, (14), p. 10)

To understand what Butler means by conscience's authority, we will study it first of all in relation to the other motivating elements of human nature. Butler's division between the passions and rational principles of action is classic to the 18th century. His originality is in considering self-love and

benevolence both as rational principles with motivational force, while concluding that the two bear no particular opposition to one another, but rather contribute to the harmony and unity of human nature. The passions, which blindly spur us to action, set pressing, incoherent and often conflicting demands which are resolved by the rational principles. While the latter help us to organize our passions and enable them to function as part of a cohesive system, only conscience ultimately draws the line between right and wrong.

Butler presents conscience as the moral faculty par excellence, informing us of our duty while motivating us to act according to its dictates. He names it interchangeably moral sense, moral reason or reflection, and divine reason, thus avoiding to side with the moral sense theorists, rational intuitionists or theologians. Instead, he expresses the paradox that conscience is a "sentiment of the understanding" or "perception of the heart" or, more appropriately, both at once (*Dissertation*, (1), p. 247).

Butler combines elements of Lord Shaftesbury's[7] reflex approbation, Samuel Clarke's moral fitness and Francis Hutcheson's moral sense, insisting on the authority of conscience to obtain a moral faculty which has the following characteristics: (i) it is reflective in that it passes judgment on our actions, character and intentions as well as on those of others; (ii) it is intuitive as to certain universal standards of duty, such as justice, veracity, prudence and a regard for the common good; (iii) it expresses moral approval and disapproval as to right and wrong: we know and feel what we should do; and (iv) it both motivates and obligates us to act accordingly.

Furthermore, according to Butler, unless we are led astray by superstition, self-partiality or self-deceit (which if reinforced by habit can lay conscience asleep and eventually lead to a corrupt conscience), our conscience,

---

[6] See Millar, in Cunliffe 1992, pp. 298-299 and Darwall 1995, p. 261.
[7] Antony Ashley Cooper, third Earl of Shaftesbury (1671-1713).

though fallible, is reliable. It has "natural" supremacy and authority, if we consider, first of all, its function in human nature in view of the purpose for which we were intended, *i.e.* virtue; and secondly, the role it plays in our constitution, that of guiding all other principles which are subordinate to it.

In Butler's view, contrary to an inanimate object or a sentient but not human being, we are agents, responsible for our flourishing. It is our conscience which, in claiming authority, reveals to us our duty and makes us morally accountable for our acts. In stating that conscience is the very faculty which makes moral agency possible, Butler suggests a third argument for the authority of conscience, which we will call the "autonomist" line of argument, following Darwall and Jerome B. Schneewind[8].

Finally, we will discuss all three lines of argument in our concluding chapter, with a view to determining whether Butler's ethics holds its ground without the theistic teleology which serves as its foundation. Firstly, let us turn to Butler's account of human nature.

---

[8] See Darwall 1995, p. 275 *ss.* and J. B. Schneewind 1998.

CHAPTER 1

Butler's Moral Psychology

As we outlined in our introduction, Butler's view of human nature treats the self as a unit, composed of several parts, each in special relation to the other. Furthermore, he orders the various principles in human nature according to a hierarchy of levels. The role of the reflective principles of self-love and benevolence is to adjust the particular passions, affections and appetites in attaining their respective ends, the good of the self and the good of others. The faculty of moral conscience crowns all other principles, applying reasonable self-love and benevolence in cases of morality and issuing overriding reasons to act morally. We will study each of these components separately, then the relation they bear to one another. Let us examine first of all the role of the particular passions, affections and appetites according to Butler.

1.1 The Particular Passions

For Joseph Butler, the system of human nature includes particular passions, affections and appetites, which he also calls instincts, desires, inclinations or principles of action. As Charlie Dunbar Broad says: "These are what we should call impulses to or aversions from particular kinds of objects." (1930, p. 60). In Butler's own words:

> 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. (*Sermons*, The Preface, (18), p. 12)

The appetites Butler mentions are hunger, thirst and the sensual appetites, whereas he uses the terms "passions" and "affections" more or less

interchangeably. Austin Duncan-Jones notes that Butler appears unclear in his terminology:

> Butler does not explain how he distinguishes a "passion", an "appetite", and an "affection" from one another. Sometimes he uses all these words together, or two of them, to stand for a certain general class of human motives. They do not seem to be quite interchangeable. [...] But it is not part of his purpose to analyse the psychological distinctions which underlie these varied names. In his view, all the passions, affections, and appetites occupy the same rank in man's constitution. (1952, p. 45)

Butler gives us many different examples of passions and affections, some tending to the private good and some to the public good. In the first category, Butler mentions envy, hatred, anger, revenge, retaliation, fear, malice, resentment, ambition and pride. In the second category, we find forgiveness of injuries, kindness, sorrow in the distress of others, tenderness, compassion, pity, rejoicing in the joy of others, mercy, friendship, love of others, liberality and charity. However, Butler nowhere maintains a distinct division between the private and the public passions. He considers the two to be complementary, enabling human beings to achieve a balance between competing parts of their nature. Moreover, many public passions, such as desire of esteem from others, contempt and esteem of others, love of society and indignation against successful vice, contribute as well to the private good, and *vice versa*:

> *Secondly*, This will further appear from observing that the *several passions* and *affections*, which are distinct both from benevolence and self-love, do in general contribute and lead us to *public* good as really as to *private*. It might be thought too minute and particular, and would carry us too great a length, to distinguish between and compare together the several passions or appetites distinct from benevolence, whose primary use and intention is the security and good of society; and the passions distinct from self-love, whose primary intention and design is the security and good of the

> individual. It is enough to the present argument, that desire of esteem from others, contempt and esteem of them, love of society as distinct from affection to the good of it, indignation against successful vice, that these are public affections or passions; have an immediate respect to others, naturally lead us to regulate our behaviour in such a manner as will be of service to our fellow-creatures. If any or all of these may be considered likewise as private affections, as tending to private good; this does not hinder them from being public affections too, or destroy the good influence of them upon society and their tendency to public good. It may be added, that as persons without any conviction from reason of the desirableness of life, would yet of course preserve it merely from the appetite of hunger; so by acting merely from regard (suppose) to reputation, without any consideration of the good of others, men often contribute to public good. (Sermon I, (7), pp. 36-38)

It is important to emphasize that Butler does not draw a fine line between public and private passions, for most of the private passions he lists seem to contain negative sentiment towards others, while the public passions closely resemble virtues. Butler himself insists that he has no bias concerning the passions and that both are necessary and complementary in human nature.

According to Butler, the passions, affections, and appetites are "particular" in that they rest in external objects as their ends, whereas the principles or affections[9] of self-love and benevolence are "general" in that they aim at bringing about internal states of the individual or of the collectivity, *i.e.* happiness or well-being. For example, there is the particular affection of compassion towards others which prompts us to do good in specific instances and then there is the general rational principle or affection of benevolence

---

[9] Butler uses the terms 'principle' and 'affection' concerning self-love, benevolence and the passions, whereas he dubs the first two *general* principles or affections, and the latter *particular* passions or affections. See Sermon I, (7), footnote 1, at p. 36: "The former of these actions is plainly to be imputed to some <u>particular</u> passion or <u>affection</u>, the latter as plainly to the <u>general affection</u> or principle of self-love." (our emphasis)

which is a desire for the global welfare of humanity and which uses the passions in order to attempt to obtain this common good. Like self-love, which aims at the happiness of the individual, benevolence aims at happiness overall, rather than at a particular external object. As Wayne G. Johnson explains:

> Butler contends that there is an important distinction between the various desires which human beings possess. There is, first of all, a class of "primary appetites" or first order desires made up of particular desires, such as our desire for food, shelter, sex. Each of these involve a particular external object which, as we say, we desire. Also included among these first order desires are benevolent impulses such as the desire to help an injured child or to entertain a friend, and malicious impulses such as the desire to blacken someone's eye. Such impulses or desires are part of our being. According to Butler, these first order desires must be distinguished from "self-love", a second order desire, which is our desire for happiness, satisfaction, etc. (1992, p. 255)

As we will see in Section 1.3, the general affections of benevolence and self-love are rational principles according to Butler since they employ reason in organizing the particular passions and affections in order to attain their ultimate ends. In contrast, the passions aim at singular objects, for example, in the case of hunger, food, in the case of ambition, success, in the case of desire of esteem, good reputation:

> Hunger is to be considered as a private appetite; because the end for which it was given us is the preservation of the individual. Desire of esteem is a public passion; because the end for which it was given us is to regulate our behaviour towards society. The respect which this has to private good is as remote as the respect that has to public good: and the appetite is no more self-love, than the passion is benevolence. The object and end of the former is merely food; the object and end of the latter is merely esteem: but the latter can no more be gratified, without contributing to the good of society; than the former can be gratified, without

> contributing to the preservation of the individual. (Sermon I, (7), p. 37, footnote 1)

Desire of esteem is no more benevolence than hunger is self-love.

Some authors question Butler's particular/external *versus* general/internal distinction between the passions and the principles of self-love and benevolence. According to C.D. Broad, Butler is mistaken in stating that the passions never aim at an internal state of the individual and rest only in external objects as their end. He illustrates this with the example of hunger, whose object is not food, but to eat food and thus relieve the impulse of hunger: "In fact the object of an impulse is never, strictly speaking, a thing or person; it is always to change or to preserve some state of a thing or person." (1930, p. 67).[10] Duncan-Jones provides a solution to this dilemma with which we agree, that of replacing Butler's use of the word "object" by "objective":

> Butler seems here to be accepting uncritically the colloquial use of the word "object", when it is combined with words standing for desires or purposive actions. It will not always be possible to follow his usage, and when something less elliptical is needed the word 'objective' will be used. An objective is the state of affairs which the passion, if unimpeded, tends to bring about. (1952, p. 49)[11]

Although Duncan-Jones concurs with Broad's assessment that particular passions can have internal states of the individual as their aim as well as external objects, he specifies that such an internal state would not be general, *i.e.* the agent's overall happiness, as in the case of self-love:

> Butler gives no analysis of the meaning of "external" and "internal", or of "object", or of the relation between

---

[10] See also Reginald Jackson 1943, p. 128: "Yet it is obvious that the objects of many affections besides self-love, including many 'appetites of sense', are states of the owners of the affections."

[11] See also Henson 1988, p. 34: "Butler does not discuss the ontological category of objects of desire; he leaves us free to think of them as states of affairs".

> a passion and an object. His language suggests that an object of someone's passion is internal if, and only if, it consists of or includes a state of feeling on the part of that person; and that otherwise it is external. Yet there must surely be passions whose objects are, in this sense, internal; for example, the desire to get rid of a persistent worry, or the appetite for thrills of various kinds. These are, in Butler's language, particular, and distinct from self-love. [...] We may concede to Butler that the objectives of self-love are internal in the sense explained, but we cannot deny that the objectives of particular passions may be internal also. The distinction between the passions and self-love will have to be found in the fact that the former are "particular", and self-love is "general". (1952, pp. 48-50)

We must remember that Butler wishes to emphasize that the object of self-love is the individual's happiness on the whole, and that it only pursues a particular passion as a means to this end. The same applies to benevolence, which employs rational instrumentality to achieve the common good. Although it is difficult to determine whether an action is motivated by a particular passion or by self-love, since they often coexist, the two remain distinct in Butler's view:

> Self-love and any particular passion may be joined together; and from this complication, it becomes impossible in numberless instances to determine precisely, how far an action, perhaps even of one's own, has for its principle general self-love, or some particular passion. But this need create no confusion in the ideas themselves of self-love and particular passions. (*Sermons*, The Preface, (36), p. 21)

We will now illustrate the role of the particular passions in Butler's theory of human nature with the examples of compassion and resentment (along with forgiveness of injuries), to which he devotes four of his *Fifteen Sermons*.

1.2 The Role of the Particular Passions in Human Nature – The Examples of Compassion and Resentment

In a footnote at the beginning of Sermon V, *Upon Compassion*, Butler contests Thomas Hobbes' definition of pity as the "imagination or fiction of future calamity to ourselves, proceeding from the sense [...] of another man's calamity" ((1), p. 84). According to Butler, this definition equates compassion with fear, which is another affection altogether. He goes on to describe three different feelings arising from the sight of another in distress: the first being the equivalent of sympathy, the second a sense of relief that we are not the one suffering, and the third an apprehension that it could be our turn next:

> There are often three distinct perceptions or inward feelings upon sight of persons in distress: real sorrow and concern for the misery of our fellow-creatures; some degree of satisfaction from a consciousness of our freedom from that misery; and, as the mind passes on from one thing to another, it is not unnatural from such an occasion to reflect upon our own liableness to the same or other calamities. (Sermon V, (1), footnote 1, p. 84)

However, says Butler, only the first of these three perceptions is properly the affection of compassion, which has at its object the person in need and which impels us to come to her aid. A compassionate person is not the same as a cowardly person, and when we say that we have more compassion towards those closer to us, we do not mean that we fear our friends more than strangers. Butler considers that Hobbes makes the "philosopher's mistake" of substituting one thing for another, *i.e.* fear for compassion, in the name of defending the theory that all actions are necessarily selfish (Sermon V, (15), p. 95). Butler believes that morality should concord with "plain common sense"[12] (or at least

[12] This may be true, but as Broad reminds us, "although common-sense here happens to be right and the philosopher to be wrong, [...] this is no reason to prefer common-sense to philosophy." As philosophers, we must explain why Hobbes is mistaken. (Broad 1930, p. 65)

not contradict it), which forces us to recognize that there are "public" affections or passions in human nature which aim directly at increasing the happiness of others and relieving their misery.

As proof of goodwill among men, Butler submits that we rejoice at the prosperity of others and at the possibility of being able to contribute to it. He then notes that we are grieved at the distress of others to an even greater extent and wish to diminish their pain. The reason why our desire to relieve others' suffering is stronger than our joy at their good fortune, explains Butler, is that the former targets a specific need of assistance which is answered by the affection of compassion.

Along Hobbesian lines, Butler's reasoning could be challenged with the retort that joy at others' happiness is weak since tainted with envy, and that compassion towards others is expressed with a view to ensuring that they will come to our assistance when we are in need. In a similar vein, Elliott Sober and David Sloan Wilson discuss whether "empathy and sympathy are able to evoke altruistic desires because people don't like experiencing these emotions and therefore wish to do what they can to extinguish them." (1998, p. 232). The results of different psychological studies which they review, however, indicate that participants wish to receive confirmation that others' suffering has been alleviated, rather than to remain uninformed of the outcome. Moreover, participants are content when the sufferer is helped by someone other than themselves[13]:

> Empathy and sympathy are emotions. When they occur, do they trigger altruistic desires? Common sense suggests that they do; empathy and sympathy sometimes elicit helping behavior, and it makes sense to see this behavior as tracing back to the desire to

[13] See also Duncan-Jones 1952 at p. 110: "We find, for instance, that a man often seems to desire some advantage for someone else, and to be pleased when he obtains it, although it was not in his own power to confer it."

> improve the other person's situation. The causal chain seems to be this:
>
> Emotions of empathy and sympathy → Desire to help → Helping
>
> (Sober and Wilson 1998, pp. 231-232)

Sober and Wilson remark that people are more easily moved by the suffering of those near and dear, than by calamity affecting those further removed, such as foreigners in a far-away land. They suggest that a distant calamity fails to elicit an emotional response in us. Along Butlerian lines, we submit instead that this may be due to our sense of powerlessness to remedy the effect of the calamity, since, as note Sober and Wilson, compassion still prompts the other-regarding desire to help in such a case. Furthermore, our attachment to those closer to us is deeper, and we are touched by their pain to a greater degree.

Butler reminds us that it is more readily in our power to contribute to the happiness of others by refraining from doing them harm or by relieving their misery, than by positively contributing to their happiness, and that this is the principal function which compassion fulfills in human nature:

> The social nature of man, and general good-will to his species, equally prevent him from doing evil, incline him to relieve the distressed, and to promote the positive happiness of his fellow-creatures: but compassion only restrains from the first, and carries him to the second; it hath nothing to do with the third.
>
> The final causes then of compassion are to prevent and to relieve misery. (Sermon VI, *Upon Compassion*, (2), pp. 98-99)

Butler's teleological argument surfaces at this point. He explains that we are sentient beings capable of joy and happiness, as well as of pain and

suffering. The constitution of our nature being so designed, and it being significantly easier to do others harm than to do them good, we are under a special moral obligation to avoid increasing their misery and to relieve it when we can. If we study compassion, we will observe that its role is to restrain envy, resentment, unreasonable self-love and other "principles from which men do evil to one another" when ungoverned (Sermon VI, (4), p. 99). None of these principles has harm in itself as its end, but each serves as a counterbalance to the other principles in order to ensure unity and harmony in human nature. As explains Broad:

> According to [Butler] none of these [principles] is intrinsically evil. Wrong-doing is always the excessive or inappropriate functioning of some principle of action which is right when acting in its due degree and in its proper place. It is like a watch with a spring which is too strong for its balance-wheel, or a constitution in which one of the estates of the realm usurps the functions of another. (1930, p. 56)[14]

How then do we account for vicious actions, whereby human beings inflict harm upon one another? How do we explain the passion of hatred? Butler's answer is that hatred serves as a check to love and *vice versa*, the one preventing the other from occupying a "disproportionate" or "unnatural" space in our economy, and the two together ensuring the welfare of the individual and of others. Just as men will gratify a passion to the extent of inflicting injury on others, they will indulge in a passion at their own expense, sometimes to the point of ruin. However, according to Butler, this is only proof that the true end of the passion was thwarted, and that it no longer serves either self-love or benevolence:

---

[14] See also A. Lefevre 1899, p. 138: "As there is no general rational principle of self-hatred, so neither is there any general rational principle of malevolence toward our fellows. Further, particular affections never make for evil for its own sake".

> [...] mankind have ungoverned passions which they will gratify at any rate, as well to the injury of others, as in contradiction to known private interest: but that as there is no such thing as self-hatred, so neither is there any such thing as ill-will in one man towards another, emulation and resentment being away; whereas there is plainly benevolence or good-will: there is no such thing as love of injustice, oppression, treachery, ingratitude; but only eager desires after such and such external goods: which, according to a very ancient observation, the most abandoned would choose to obtain by innocent means, if they were as easy, and as effectual to their end: [...] and that the principles and passions in the mind of man, which are distinct both from self-love and benevolence, primarily and most directly lead to right behaviour with regard to others as well as himself, and only secondarily and accidentally to what is evil. (Sermon I, (12), pp. 42-43)[15]

For example, if compassion were truly delight in power over another as Hobbes claims, it could take an evil form (such as relishing their misfortune), rather than the desire to relieve another's suffering, which is a motive to deliberately do good to another. Furthermore, when we feel that the other's suffering was brought about by her own wrongdoing, we do not feel compassion, as in the case of a criminal in distress. This demonstrates that the end of compassion is to prevent mischief, and not contribute to it.

In opposition to the Stoics, Butler does not view compassion as a weakness, but as a necessary guide to human conduct: when possessed and exercised in due proportion, it promotes both personal and the general happiness. We reap satisfaction in helping relieve another's misery, in knowing that we succeeded in doing so and that her suffering is alleviated or diminished; or simply in knowing that we did what we could to provide assistance, even

---

[15] See also Lefevre 1899, p. 141: "So for Butler, evil action results from permitting our desires to run to wild extremes, unguided and ungoverned by the supreme faculty of our nature which claims sovereignty; it is the consequence of the destruction of the due and just proportion prescribed by reason."

when our efforts fail. We also experience relief in the belief that another would do the same for us if we were one day in a similar predicament. As to the person in distress, she feels her pain subside and (or at the very least) is comforted in discovering that she is not alone in her suffering but has received succor from others and may do so again in the future.

Butler was well aware that compassion supplements reason and moral principles in preventing us from doing harm to others and in motivating us to assist them when in need. As note Sober and Wilson, "It is possible that other-directed desires come into existence without the mediation of an empathic pathway" (1998, p. 237); but empathic and sympathetic emotions greatly aid and abet reason in carrying forth its dictates. The rational principle of benevolence would be less effectual if it did not have the assistance of affections such as compassion, just as self-love would be less effectual if it did not have hunger, thirst and the many passions setting about to obtain "objects" to further its end, personal happiness:

> Is it possible any can in earnest think, that a public spirit, *i.e.* a settled reasonable principle of benevolence to mankind, is so prevalent and strong in the species, as that we may venture to throw off the under affections, which are its assistants, carry it forward and mark out particular courses for it; family, friends, neighbourhood, the distressed, our country? The common joys and the common sorrows, which belong to these relations and circumstances, are as plainly useful to society; as the pain and pleasure belonging to hunger, thirst and weariness are of service to the individual. In defect of that higher principle of reason, compassion is often the only way by which the indigent can have access to us [...]. (Sermon V, (10), p. 92)

In the same way as compassion encourages benevolence, let us now examine how resentment contributes to self-love. Just as compassion is a restraint on the motive of injury, *i.e.* resentment, resentment is a check on

compassion. Compassion in excess can prevent its owner from protecting herself against injury. Resentment ensures that the compassionate are not too lenient towards wrong-doing. Butler divides resentment into two categories: hasty and sudden, or settled and deliberate resentment. While sudden resentment resembles anger and has as its motive protection from harm independently of any idea of wrong-doing, deliberate resentment aims at countering injustice, oppression and malice and entails a battle against moral evil (Sermon VIII, *Upon Resentment*, (3)-(6), pp. 123-125):

> From hence it appears, that it is not natural, but moral evil; it is not suffering, but injury, which raises that anger or resentment, which is of any continuance. The natural object of it is not one, who appears to the suffering person to have been only the innocent occasion of his pain or loss; but one, who has been in a moral sense injurious either to ourselves or others. This is abundantly confirmed by observing what it is which heightens or lessens resentment; namely, the same which aggravates or lessens the fault: friendship and former obligations, on the one hand; or inadvertency, strong temptations and mistake on the other. (Sermon VIII, (6), p. 126)

The natural function of resentment, however, can be misused. In the case of sudden resentment, Butler names the abuse of resentment in strong natures *passion* (as in a person subject to fits of temper), and in weak natures *peevishness* (as in those who continually find cause for resentment, even without foundation). In the case of deliberate resentment, the many abuses are the imagination of injuries where there are none, the exaggeration of same, resentment without demerit as against the innocent, a reaction of indignation out of proportion to the harm caused, and the desire to inflict injury simply to gratify a feeling of resentment, where ill-founded. These abuses are habitually accompanied by self-partiality and a refusal to listen to reason (Sermon VIII, (9)-(10), pp. 128-129).

As is the case with compassion, the function of resentment, according to Butler, is to supplement a rational principle, *i.e.* self-love. Firstly, the passion prevents human beings from inflicting harm on others needlessly, through fear of their resentment. Secondly, it ensures the punishment of wrong-doing, through moral indignation:

> The good influence which this passion has in fact upon the affairs of the world, is obvious to every one's notice. Men are plainly restrained from injuring their fellow-creatures by fear of their resentment; and it is very happy that they are so, when they would not be restrained by a principle of virtue. And after an injury is done, and there is a necessity that the offender should be brought to justice; the cool consideration of reason, that the security and peace of society requires examples of justice should be made, might indeed be sufficient to procure laws to be enacted, and sentence passed: but is it that cool reflection in the injured person, which, for the most part, brings the offender to justice? Or is it not resentment and indignation against the injury and the author of it? (Sermon VIII, (13), p. 131)

Thus resentment, although it immediately and directly aims at inflicting harm, is justified because its ultimate aim is to prevent injury to oneself and to others. Furthermore, for this reason, it is necessary in the constitution of human nature. Henry Sidgwick explains Butler's theory:

> Others, however, think that a deliberate and sustained desire to punish wrong-doers is required in the interests of society, since the mere desire to realise Justice will not practically be strong enough to repress offences: and that it is as serious a mistake to attempt to substitute the desire of Justice for natural resentment as it would be to substitute prudence for natural appetite in eating and drinking, or mere dutifulness for filial affection. (1981 (1907 7th ed.), p. 323)

In addition, forgiveness of injuries serves as a check to resentment. Since the final aim of resentment is to prevent injury and disorder, it should not

be indulged for its own sake if it is to serve its main purpose, the private and the common good. Unfortunately, "malice begets malice" (Sermon IX, *Upon Forgiveness of Injuries*, (5), p. 137), and unbridled resentment, by its very nature, would encourage a string of injuries and injustices. This is why retaliation cannot be made a universal law of conduct if it is to ensure the general welfare:

> Every natural appetite, passion and affection may be gratified in particular instances, without being subservient to the particular chief end, for which these several principles were respectively implanted in our nature. And, if neither this end, nor any other moral obligation be contradicted, such gratification is innocent. [...] But the gratification of resentment, if it be not conducive to the end for which it was given us, must necessarily contradict, not only the general obligation to benevolence, but likewise that particular end itself. The end, for which it was given, is to prevent or remedy injury, *i.e.* the misery occasioned by injury; *i.e.* misery itself: and the gratification of it consists in producing misery; *i.e.* in contradicting the end for which it was implanted in our nature. (Sermon IX, (9), p. 139-140)

Therefore, resentment is restrained by forgiveness of injuries, which itself stems from compassion.

As with compassion, Butler's description of the role and function of resentment in the system of human nature reflects a naturalistic teleology. When resentment no longer serves its purpose, that of preventing injury, but contradicts it in inflicting injury for its own sake, out of retaliation and revenge, it is "unnatural" and an abuse of the passion. Butler argues that resentment is in no way inconsistent with goodwill, proof being that we often harbour greater grudges against those who are closest to us and for whom we care the most. The passion of resentment, as that of compassion, functions properly when it serves both self-love and benevolence:

> Resentment is not inconsistent with good-will: for we often see both together in very high degrees; not only in parents towards their children, but in cases of friendship and dependence, where there is no natural relation. These contrary passions, though they may lessen, do not necessarily destroy each other. We may therefore love our enemy, and yet have resentment against him for his injurious behaviour towards us. But when this resentment entirely destroys our natural benevolence towards him, it is excessive, and becomes malice or revenge. The command to prevent its having this effect, *i.e.* to forgive injuries, is the same as to love our enemies; because that love is always supposed, unless destroyed by resentment. (Sermon IX, (13), p. 141)

Thus, the end of forgiveness of injury is to curb abuses of the passion of resentment by helping it resume the initial function of all passions, that of furthering rational love: of oneself, of one's neighbour, even of one's enemies[16]. To Sidgwick, this resembles Utilitarian reasoning:

> For the Christian code is widely thought to prescribe a complete and absolute forgiveness of such offences, and many Christians have endeavoured to carry out this rule by dismissing the offences as far as possible from their minds, or at least allowing the memory of them to have no effect on their outward conduct. [...] If we ask, therefore, how far forgiveness is practically possible, the answer seems admittedly to depend on two considerations: (1) how far the punishment to which resentment prompts is really required in the interests of society, and (2) how far, if so, it will be adequately inflicted if the person wronged refrains from inflicting it. But, obviously, so far as we allow the question to be settled by these considerations we are introducing a method difficult to distinguish from the Utilitarian. (1981, p. 322)

However, for Butler, our duty of goodwill to all men does not stem only from social convention or moral code. Given our sentient nature, we are

subject to pleasure and pain, and this itself creates a duty not to inflict unnecessary suffering on (ourselves or) others:

> It is not man's being a social creature, much less his being a moral agent, from whence *alone* our obligations to good-will towards him arise. There is an obligation to it prior to either of these, arising from his being a sensible creature; that is, capable of happiness or misery. Now this obligation cannot be superseded by his moral character. What justifies public executions is, not that the guilt or demerit of the criminal dispenses with the obligation of good-will, neither would this justify any severity; but, that his life is inconsistent with the quiet and happiness of the world: that is, a general and more enlarged obligation necessarily destroys a particular and more confined one of the same kind inconsistent with it. Guilt or injury then does not dispense with, or supersede the duty of, love and good-will. (Sermon IX, (15), p. 142)

We will now explore further how the rational principles of self-love and benevolence employ the passions in carrying out their mission.

1.3 The Rational Principles of Self-Love and Benevolence

a) Self-Love

Butler distinguishes between what he calls "cool" or "reasonable" self-love and "supposed" self-interest. The first is part of our rational nature and is an attempt to ensure our long-term interest or happiness. The second usually refers to the ungoverned indulgence in the passions, irrespective of rationality (*i.e.* the principles of self-love, benevolence and conscience). In the *Dissertation*, Butler equates reasonable self-love with the virtue of prudence, and imprudence with "folly". As elaborates Duncan-Jones:

---

[16] Paul A. Newberry 2001 interprets Butler's definition of forgiveness as the checking of revenge, or forbearance, rather than the overcoming of resentment.

> Butler's most typical statements about self-love concern what he also calls "cool self-love", "cool and reasonable concern" for oneself, "reasonable self-love", "cool consideration that" an action "will be to my own advantage", "general desire of happiness", "manifest and real interest", and so on. It "belongs to man as a reasonable creature, reflecting upon his own interest or happiness" (S. 11.5). He sometimes contrasts self-love, in this most typical sense, with "supposed interest", or "supposed self-love", that is, with people's false notions of what will bring them happiness: sometimes also – and this is harder to reconcile, as we shall see, with his general account of self-love – with "immoderate self-love" (S. 11.9), or "unreasonable and too great regard to ourselves", "over-fondness for ourselves" (S. 10.6). (1952, pp. 59-60)

Butler believes that most of us possess self-regard to an extent which permits us to effectively secure our private good. Moreover, since there exist private passions which are of exclusive concern to the individual and do not regard others, we are morally entrusted with ourselves in a special manner (Sermon XII, (17), p. 194). In the same way as we are conscious of our own existence, we are more keenly aware of our interests, even though we feel equal affection for others. For Butler, endeavouring to pursue our own happiness is thus both an actual possibility and a duty.

However, in Butler's view, our true happiness lies in attributing "due proportion" to the parts which make up our nature and employing them towards the purpose they were meant to serve. For example, an excessive passion or inordinate self-regard would thwart its proper end and result in anxiety or misery. The notion of "due proportion" concerns not only the degree to which self-love or a passion prevails, but also its relative weight vis-à-vis the other principles. Contrary to brutes, we have the capacity to regulate our behaviour through reason, rather than blindly following passion, appetite or instinct as they arise.

Butler's naturalistic approach to morals considers that if we study our makeup or constitution, we will see that certain principles are superior to others in kind or in nature. Even though, in practice, self-love and benevolence are weaker than the passions and more difficult to follow, they assert themselves as superior. Butler distinguishes "mere power" from the "authority" of a principle[17], which he illustrates with an analogy to civil government: in a corrupt state, where the legislature or the judiciary has authority under the constitution, its power may still be usurped:

> 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. (Sermon II, (11), p. 55)

The role of self-love as a principle of reflection in Butler's ethical theory can be better understood if we apply it in relation to his treatment of self-deceit: "Though a man hath the best eyes in the world, he cannot see any way but that which he turns them." (Sermon X, *Upon Self-Deceit*, (4), p. 153); "It is as easy to close the eyes of the mind, as those of the body" (Sermon X, (11), p. 159). A human being engrossed by vanity, blinded by self-deceit, mesmerized by a vision of her own happiness and totally preoccupied with her interest, will neglect many of the joys and satisfactions of life. In Butler's opinion, she will lose sight of the purpose of her actions due to her obsession with her own agency. For Butler, we avoid self-deceit through detached

---

[17] See Sermon II, (14), p. 57: "All this is no more than the distinction, which everybody is acquainted with, between *mere power* and *authority*: only instead of being intended to express the difference between what is possible, and what is lawful in civil government; here it has been shown applicable to the several principles in the mind of man."

reflection concerning the long-term efficacy of our present and past actions in procuring our happiness and furthering our welfare. Self-deceit is an unfairness of mind towards the self, as malicious behaviour is an unfairness of mind *vis-à-vis* others. It is precisely for this reason that self-deceit and vice are inextricably linked according to Butler. Absence of self-distrust is due to a lack of reflection, and "supposed self-interest" prevents us from remaining at a critical distance from ourselves. To avoid self-deceit, we need to judge our conduct as if it were another's, from an objective standpoint. With a less biased view of our nature, we have a greater chance of furthering our own happiness or interest through self-love. Butler believes that we are better off with passions ungoverned by self-love, than with an exaggerated sense of self-love. Such a "deep and calm source of delusion" not only disappoints itself and contradicts its own end, private good, but "undermines the whole principle of good; [...] and corrupts conscience, which is the guide of life." (Sermon X, (16), p. 163).

We will complete our discussion of Butler's concept of self-love in Chapter 2.

b) Benevolence

In Sermon I, *Upon Human Nature*, Butler describes benevolence as a natural principle which serves a purpose for society –the good of others or the public good – analogous to the purpose self-love serves for the individual – private good or interest. In Sermons XI and XII, *Upon the Love of our Neighbour*, he qualifies both self-love and benevolence as "general" principles or affections. However, while commentators agree that for Butler, self-love is a rational principle, some treat Butler's benevolence as a particular affection. We concur with the Broad-Taylor[18] view (often dubbed the "traditional

[18] See Broad 1930, A.E. Taylor 1926, Grave 1952, Roberts 1973, Penelhum 1985 (for Penelhum, benevolence is a generic term covering all other-regarding affections), Riddle 1959 and Raphael 1949 (Riddle and Raphael treat Butler's benevolence as "both a rational principle applying to humanity in general and a particular affection applying in specific

interpretation") that self-love and benevolence are on an equal footing as rational principles in Butler's moral theory, given his similar treatment of the two:

> I have assumed throughout that [Butler] regards benevolence as a general principle which impels us to maximise the happiness of humanity without regard to persons, just as he certainly regards self-love as a general principle leading us to maximise our own total happiness. I think that this is what he does mean. But he sometimes tends to drop benevolence, as a general principle co-ordinate with self-love, rather out of sight, and to talk of it as if it were just one of the particular impulses. (Broad 1930, p. 71)

The commentators who maintain the opposing view insist that Butler groups benevolence, when comparing it to self-love, with the "other" particular passions. It is important to note that, in the passages usually quoted by these philosophers, Butler's intention is to illustrate that benevolence relates to self-love in the same way as any "other" particular passion or affection, *i.e.* that it is as "interested" or as "disinterested" as they are. He is intent on demonstrating that there is no specific contradiction between self-love and benevolence, any more than between self-love and the particular passions, and that only self-love is truly "interested":

> But whatever occasioned the mistake, I hope it has been fully proved to be one; as it has been proved, that there is no peculiar rivalship or competition between self-love and benevolence; that as there may be a competition between these two, so there may also between <u>any particular affection whatever</u> and self-love; that <u>every particular affection, benevolence among the rest, is subservient to self-love</u> by being the instrument of private enjoyment; and that in one respect benevolence contributes more to private interest, *i.e.* enjoyment or satisfaction, than any other of the

---

instances" (Riddle 1959, p. 361)) *versus* Sidgwick 1981, Jackson 1943, McPherson 1948 & 1949, Duncan-Jones 1952, and Frey and McNaughton in Cunliffe 1992.

> particular common affections, as it is in a degree its own gratification. (our emphasis) (Sermon XI, (19), p. 181)

Although the end of the above passage, by its wording, leads us to believe that Butler does not put benevolence on the same level as self-love, he clearly states at the beginning that his goal is quite the opposite. We must understand the term "affections" to indicate, at times, all the principles of human nature, self-love and benevolence included. The following passage illustrates the parity of self-love and benevolence for Butler:

> Further, the whole system, as I may speak, of affections (including rationality) which constitute the heart, as this word is used in Scripture and on moral subjects, are each and all of them stronger in some than in others. Now the proportion which the two general affections, benevolence and self-love, bear to each other, according to this interpretation of the text, denominates men's character as to virtue. (our emphasis) (Sermon XII, (11), p. 191)

It is apparent from the above that when Butler speaks of benevolence or self-love in comparison with any "other" particular affection, he does not mean to lower either of these general principles to the rank of a particular passion.

In discussing the status of benevolence, certain writers emphasize that the passions as well use reason to achieve their ends, as in the cases of revenge and ambition, yet this does not make them rational principles as such. In response to this contention, while Butler agrees that the passions employ instrumental rationality (for example, in the case of deliberate resentment); what he calls a "rational principle" applies reason not only in choosing means to attain its ends, but also as a guide to conduct in choosing the ends themselves. For Butler, a principle is "rational" if it uses both calculative and "directive" rationality, to cite David McNaughton: there are "two quite different ways in which a principle can be rational: roughly, the distinction

between reason as a guide to help us achieve a given end, and reason as directive, determining what our ends should be."(in Cunliffe 1992, p. 280). However, McNaughton disagrees that Butler's benevolence is "rational" in the latter sense of the term:

> It is undeniable that benevolence is a rational principle in that its exercise requires calculative rationality. It is a controversial thesis, which by no means follows from that claim, that it is what we might call a principle of reason, a principle whose dictates are authoritative, in Butler's sense. Self-love is indeed a rational principle which is also a principle of reason, but we must not suppose that the one entails the other. (p. 281)

The main difficulty which causes commentators to rank benevolence below self-love is that Butler credits it with less motivational force than self-love. In our opinion, Butler clearly considers benevolence to be as fully a "rational principle" as self-love. However, neither benevolence nor self-love are "authoritative" in the way Butler understands conscience to be, although they are superior principles to the passions. McNaughton seems to equate the superiority of a principle with its authority, and its authority with rationality: "Butler's account of superiority rests, as is well known, on a difference between the felt strength and the authority of a principle of action. [...] That authority is a rational one; the verdicts of a superior principle provide better reason to act than the promptings of an inferior one." (p. 280). What McNaughton neglects to point out is that, for Butler, conscience is superior "over all others" (Sermon II, (14), p. 57). It is not only superior; it has supreme authority over all other principles in human nature.

Butler's hierarchy (leaving out conscience for the moment) can be described as follows: passions use calculative rationality and are stronger than the rational principles; self-love and benevolence use calculative and "directive" rationality, but are often weaker than the passions; benevolence, which is the sum of virtue in an ideal being but requires the guidance of

conscience in a human being, is usually weaker than self-love. For this reason, as concludes Raymond Gillespie Frey (although we disagree with the place allowed to benevolence in his interpretation of Butler), Butler insists on giving us "an interested reason to pursue benevolence":

> The position, then, is one concerned with motivation and the effective prospect of motivating actual men to be benevolent, in the face of prudence and of their powerful desire for their own happiness. (R.G. Frey, in Cunliffe 1992, p. 267)

For Butler, the motivational force of a principle, its strength or degree, is a separate issue from the nature of a principle, or its kind. While the strength of a principle could ultimately affect the possibility of a moral claim, by rendering an "ought" ineffectual in practice, "Everything is what it is, and not another thing." (*Sermons*, The Preface, (39), p. 23). Like self-love, benevolence is directed by reason. Since the good of others consists in the enjoyment of those objects which are best suited to bring them happiness, benevolence governs the public passions by 1) helping them to choose appropriate objects for their fulfilment and 2) preventing them from having free reign and countering their initial aim through excess or deprivation:

> Thus, when benevolence is said to be the sum of virtue, it is not spoken of as a blind propension, but as a principle in reasonable creatures, and so to be directed by their reason: for reason and reflection comes into our notion of a moral agent. And that will lead us to consider distant consequences, as well as the immediate tendency of an action: it will teach us, that the care of some persons, suppose children and families, is particularly committed to our charge by Nature and Providence; as also that there are other circumstances, suppose friendship or former obligations, which require that we do good to some, preferably to others. Reason, considered merely as subservient to benevolence, as assisting to produce the greatest good, will teach us to have particular regard to these relations and circumstances; because it is plainly for the good of the

> world that they should be regarded. (Sermon XII, (27), pp. 197-198)

McNaughton comments that benevolence cannot be a "general" principle in the same way as self-love, which employs all of our passions in attaining our overall interest over time, since another's interest can only be a particular, and not a general, object for an agent (in Cunliffe 1992, p. 275). Broad's reply to this objection is that benevolence is a "general" principle because it concerns the happiness of others "without regard to persons" (Broad 1930, p. 71). However, McNaughton is of the opinion that Butler's benevolence is not intended to encompass humanity in its entirety since he often insists on filial and parental ties and affections (p. 271). Here we agree with Frey that Butler's main concern is convincing us to the practice of virtue, and since Butler's is a naturalistic approach which involves moral psychology, he is careful to preach only within the realm of human possibilities.[19] Butler equates the principle of benevolence with the love of our *neighbour* since it "is that part of the universe, 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." (Sermon XII, (3), pp. 186-187).

"Love thy neighbour as thyself" and "do unto others as you would have them do unto you"[20] provide the starting point for Butler's concept of benevolence, which in no way involves a disregard of self-interest. Such an inordinate regard for others, *i.e.* one which ignores self-regard, would defeat its own purpose, as in the case of liberality leaving a rich person poor and no longer able to provide for herself, let alone help others. We could imagine a similar scenario with charity, forgiveness of injuries, mercy, pity, compassion,

---

[19] In this sense, he attempts to respect what Owen Flanagan calls the "Principle of Minimal Psychological Realism (PMPR)" in ethical thought: "Make sure when constructing a moral theory or projecting a moral ideal that the character, decision processing, and behavior prescribed are possible, or are perceived to be possible, for creatures likes us." (Owen Flanagan 1991, p. 32).

and so on. According to Butler, a due proportion attributed to benevolence as a rational principle in our constitution not only promotes the good of others, but is a source of pleasure to the individual, bringing her peace of mind unequalled by any other pursuit. Butler argues that benevolence is the greatest source of personal happiness.[21] Contrary to self-love, the satisfaction benevolence brings is not tainted by sorrow or fear of death. Furthermore, benevolence is an object of affection when observed in ourselves and in others. Due to its other-regarding nature, it brings joy to the agent and to others even when it does not succeed in materially improving their situation, for example, by relieving poverty or pain. The benevolent intention and ineffective act, being in themselves good, contribute to the happiness in the world regardless of their consequences:

> In case of success, surely the man of benevolence hath as great enjoyment as the man of ambition; they both equally having the end of their affections, in the same degree, tended to: but in case of disappointment, the benevolent man has clearly the advantage; since endeavouring to do good considered as a virtuous pursuit, is gratified by its own consciousness, *i.e.* is in a degree its own reward. (Sermon XI, (13), p. 175)

Butler argues that there is no particular contrariety between self-love and benevolence, and that the conflicts between the various principles in our nature are situated more often between the particular passions and affections, on the one hand, and self-love or benevolence on the other. However, he mentions that many deny the existence or even possibility of benevolent acts for the following reasons: 1) there is anti-social conduct in society, which demonstrates principles in human beings which lead them to do evil to others;

---

[20] Rom. xiii.9; and Matt. vii.12: "whatsoever ye would that men should do unto you, do ye even so unto them".

[21] "being in good humour, which is benevolence whilst it lasts, is itself the temper of satisfaction and enjoyment." (Sermon XI, (14), p. 176). As well, since benevolence or love of our neighbour is the "temper of virtue" (Sermon XII, (8), p. 188), it "includes in it all virtues" or "all that is good and worthy" (Sermon XII, (25), p. 197 and (32), pp. 200-201).

and 2) there are people without natural affection to others. Butler replies that these objections are as much proof against self-love as they are against benevolence. In response to 2), he argues that there are persons without natural affection to themselves, *i.e.* lacking reasonable self-love, who rush to certain ruin for the gratification of a particular passion. With regard to 1), as far as Butler is concerned, there are no evil principles *per se* in human nature; what we call a vicious person is someone acting unnaturally on her passions. No one seeks evil for evil's sake alone; rather, we accomplish vile acts in order to obtain external objects (such as money, power, love, etc.), thinking that these objects will bring us happiness. However, having mistaken the means to happiness for its end, we realize that we haven't obtained what we want after all, and suffer self-condemnation as well. Thus, we have indeed poorly served self-love and our own interest.

Let us now take a closer look at the relationship between self-love and benevolence in Butler's theory.

c) The Relation between Self-Love and Benevolence

For Butler, *as thyself* in the precept "love thy neighbour as thyself" typifies the relation between benevolence and self-love, in that we act conformably to our nature by allowing due proportion to the two principles. Butler states that we must first ask ourselves "what is a competent care and provision for ourselves" (Sermon XII, (14), p. 192), and that the answer to this question depends on our nature and conditions. Each person must determine this for himself, and it would be "ridiculous [...] to determine it for another" (*Ibid*). However, without deciding what is for another's good, we can treat others with equal respect and due regard, *i.e.* as we would want to be treated by them. Butler states that loving thy neighbour as thyself implies, firstly, having the same *kind* of affection towards others as towards oneself. This prevents us from forming a notion of our own private good which excludes altogether the

consideration of others' interests. Benevolence in due proportion to self-love brings us to consider ourselves as "having a real share in [our neighbour's] happiness" (Sermon XII, (7), p. 188). Secondly, Butler submits that loving thy neighbour as thyself implies having the same *degree* of regard for others as for oneself, with the reservation that such an equality of treatment can only be imperfect:

> To these things must be added, that moral obligations can extend no further than to natural possibilities. Now we have a perception of our own interests, like consciousness of our own existence, which we always carry about with us; and which, in its continuation, kind and degree, seems impossible to be felt in respect to the interests of others. (Sermon XII, (18) p. 194)

There are many passions, affections and appetites we seek to satisfy which have no regard whatsoever to others. Because of the prevalence of these uniquely self-regarding affections, we are morally entrusted with ourselves first and foremost. Thus, even if there were to be an equality of affection to ourselves and to others, "yet regards to ourselves would be more prevalent than attention to the concerns of others" (Sermon XII, (16), p. 193).

Finally, Butler regards the principles of self-love and benevolence as incentives to one another:

> I must however remind you that though benevolence and self-love are different; though the former tends most directly to public good, and the latter to private: yet they are so perfectly coincident, that the greatest satisfactions to ourselves depend upon our having benevolence in a due degree; and that self-love is one chief security of our right behaviour towards society. It may be added, that their mutual coinciding, so that we can scarce promote one without the other, is equally a proof that we were made for both. (Sermon I, (6), pp. 35-36).

Butler reminds us that a person with an extraordinarily high degree of benevolence is not necessarily virtuous, for she could also have inordinate self-regard, to the point of permitting herself to act viciously at times. Furthermore, a corrupt or evil person does not necessarily lack benevolent behaviour, but her benevolent acts, whether occasional or not, do not make her virtuous. Benevolence as a virtue is a habitual, settled disposition according to Butler. A virtuous person does not neglect her own interest, but, to the extent possible, has the same kind and degree of regard towards others as towards herself and obeys her conscience by refraining from wrongdoing.

Both self-love and benevolence require the "under affections" to reach their goals.[22] These are the particular passions or affections, which are not general rational principles or affections like self-love and benevolence. However, as notes Glenn K. Riddle, Butler's terminology at times causes confusion as to the nature of benevolence, since among the various public affections, he mentions compassion, mercy, pity, sympathy, love of another, goodwill, charity, and benevolence.[23] In all fairness to Butler, he also uses different expressions when speaking of self-love, such as "private interest", and it can sometimes be difficult to determine whether he is speaking of the rational principle or of an "under affection", although the two remain distinct. Richard G. Henson is in agreement with us on this point:

> I have said that [Butler] regards self-love as the desire for happiness: I should add that he uses several words synonymously with each of these key terms. He does not distinguish between self-love, interest, self-interest, and love of self; and the object of this variously-named desire is known not only as happiness, but as one's interest, advantage, or good, and sometimes as one's

---

[22] See quote at page 21.

[23] See Riddle 1959 at p. 357: "The uncertainty of the place of benevolence in the hierarchy is the result of an apparent inconsistency in Butler's language." See also Broad 1930 at p. 72: "I think that [Butler] makes such apparent mistakes partly because he is anxious to show that benevolence is, as such, no more contrary to self-love than is any of the particular impulses." (our emphases)

> satisfaction, enjoyment, or pleasure. Occasionally he half-grants the legitimacy of calling some *other* self-centered desires species of self-love: when he does, he calls them "sensual selfishness", as contrasted with "cool or settled selfishness" or "cool self-love". The latter is the same as self-love *sans phrase*. (1988, p. 34)

Finally, Butler points out that, just as self-love is not happiness, the rational principle of benevolence is not the general welfare.[24] Butler considers that the alleged conflict between self-love and benevolence arises from the confusion of one of the means to happiness, *i.e.* property, and happiness itself. He argues that possessing the materials of enjoyment does not constitute happiness; and that our interest is not reduced by another having a share in that happiness (Sermon XI, (19), pp. 180-181)[25]. Happiness is the enjoyment of those objects (or objectives) which are well suited to it, including the "temper" of benevolence or the consciousness of helping others. Thus, happiness will not ensue to the degree that self-love engrosses the individual. Self-love or endeavouring to obtain happiness for ourselves is aided by benevolence, just as striving to contribute to others' happiness profits from self-love:

> Happiness consists in the gratification of certain affections, appetites, passions, with objects which are by nature adapted to them. Self-love may indeed set us on work to gratify these; but happiness or enjoyment has no immediate connection with self-love, but arises from such gratification alone. Love of our neighbour is one of those affections. This, considered as a *virtuous principle*, is gratified by a consciousness of *endeavouring* to promote the good of others; but considered as a natural affection, its gratification consists in the actual accomplishment of this

---

[24] "Happiness does not consist in self-love. The desire of happiness is no more the thing itself, than the desire of riches is the possession or enjoyment of them. People may love themselves with the most entire and unbounded affection, and yet be extremely miserable." (Sermon XI, (9), p. 170).

[25] See Simon Blackburn 1998 at pp. 142-143: "Butler rightly remarks that finding a direct opposition between a principle of self-love and a principle of altruism derives from thinking of a cash transaction: if my neighbour gets the money, I do not; if I get it, he does not, and I may have to choose one outcome over the other."

> endeavour. Now indulgence or gratification of this affection, whether in that consciousness or this accomplishment, has the same respect to interest, as indulgence of any other affection; they equally proceed from or do not proceed from self-love, they equally include or equally exclude this principle. Thus it appears, that *benevolence and the pursuit of public good hath at least as great respect to self-love and the pursuit of private good, as any other particular passion, and their respective pursuits*. (Sermon XI, (16), p. 178)

In the above-quoted passage, Butler is once again intent on demonstrating that there is no particular contrariety between self-love and benevolence, be it the rational, "virtuous principle" or general affection of benevolence, or its "under affection", the particular passion.

In the next chapter, "Butler's Arguments against Psychological Egoism", we will discuss the importance of this claim to Butler's moral theory and the possibility of altruism.

CHAPTER 2

Butler's Arguments against Egoism

2.1 Butler's Refutation of Psychological Egoism

According to Duncan-Jones, "Butler's is the classic refutation" of psychological egoism, the theory of universal selfishness, or of psychological hedonism, the theory that all of our actions are motivated by the pursuit of pleasure and the avoidance of pain (1952, p. 95). Broad states that David Hume accepted and emphasized Butler's refutation of psychological egoism, and that "[a]s a psychological theory it was killed by Butler" (1930, p. 55)[26]. This appreciation of Butler's work has been contested by many a contemporary author[27], and merits attention in light of the roles Butler attributes respectively to the passions, self-love, benevolence and conscience.

Butler wishes to refute the theory of universal selfishness, which he attributes to Thomas Hobbes. According to Butler, it is against common sense to declare that all actions are selfish or done from self-interested motives. Butler criticizes Hobbes' deriving from an inescapable fact of human nature – that all voluntary actions by definition emanate from the self, *i.e.* "are done to gratify, an inclination in a man's self" (Sermon XI, (7), p. 169) – the conclusion that all actions are necessarily selfish:

> There is a strange affectation in many people of explaining away all particular affections, and

---

[26] "he killed the theory so thoroughly that he sometimes seems to the modern reader to be flogging dead horses. Still, all good fallacies go to America when they die, and rise again as the latest discoveries of the local professors. So it will always be useful to have Butler's refutation at hand." See also T.H. McPherson 1949, Part II, p. 21: "His refutation of Hobbes's psychological hedonism in the *Sermons* is one of the clearest and most convincing pieces of reasoning to be found in any work on ethics."

[27] See R. M. Stewart 1992, Jackson 1943 and M. J. Scott-Taggart 1968 at p. 16: "Butler's arguments, in substance if not in style, rarely rise above the level that can often be overheard in a pub brawl on the question of selfishness."

> representing the whole of life as nothing but one continued exercise of self-love. Hence arises that surprising confusion and perplexity in the *Epicureans* of old, *Hobbes*, the author of *Reflections, Sentences et Maximes Morales*, and this whole set of writers; the confusion of calling actions interested which are done in contradiction to the most manifest known interest, merely for the gratification of a present passion. [...] the pursuit of these external objects, so far as it proceeds from these movements (for it may proceed from self-love), is no otherwise interested, than as every action of every creature must, from the nature of the thing, be; for no one can act but from a desire, or choice, or preference of his own. (*Sermons*, The Preface, (35), pp. 20-21)

This initial argument against psychological egoism states that when the egoist insists we only act from selfish motives, he is in fact making the trivial tautological statement that all voluntary actions proceed from our own motives:

> In short, the generalisation that all my actions spring from *my* desires or motives proves to be a disguised tautology, which must be true, given the accepted meaning of the word 'action'. (Duncan-Jones 1952, p. 97).

In Sermon XI, Butler argues that not all actions initiated by an agent are self-interested and that it is important not to confuse the objects (or objectives) of the particular passions and the object of self-love. In his view, only actions proceeding from self-love are truly interested, in that they aim at the individual's happiness, whereas actions inspired by the particular passions are neither interested nor disinterested. The interested-disinterested distinction is irrelevant as far as the passions are concerned, for they seek specific external objects for their fulfilment, *i.e.* objects which are external to the self or to an overall state of that self:

> The most intelligible way of speaking of it seems to be this: that self-love, and the actions done in

> consequence of it (for these will presently appear to be the same as to this question) are interested; that particular affections towards external objects, and the actions done in consequence of those affections, are not so. But every one is at liberty to use words as he pleases. All that is here insisted upon is, that ambition, revenge, benevolence, all particular passions whatever, and the actions they produce, are equally interested or disinterested. (Sermon XI, (11), p. 174)

According to Butler, self-love is a desire for one's happiness, and that happiness consists in the enjoyment of things to which our nature is adapted, which enjoyment arises from the gratification of the passions. Each particular passion aims at an object which is external to the agent, and the object is characterised by a "prior suitableness" with the passion in question. As the passions further our private good as well as the public good, certain objects are more naturally suited to certain particular passions as means to further their ends. When these objects are used inappropriately, they thwart the passions they were meant to satisfy, with the result that the passions appear unnatural, distorted, perverse, even vicious. Just as we do not swallow stones to appease hunger, we do not normally pursue the good of others to become rich, nor hoard our riches to make others happy:

> That all particular appetites and passions are towards *external things themselves*, distinct from the *pleasure arising from them*, is manifested from hence; that there could not be this pleasure, were it not for that prior suitableness between the object and the passion: there could be no enjoyment or delight from one thing more than another, from eating food more than from swallowing a stone, if there were not an affection or appetite to one thing more than another. (our emphasis) (Sermon XI, (6), pp. 167-168)

According to Sober and Wilson, Butler's argument fails to answer the hedonist's conviction that, when we pursue external objects (objectives) which bring us satisfaction, we are ultimately motivated by the antecedent desire for

pleasure. They understand Butler as claiming that the only route to pleasure is the satisfaction of a desire for an external thing, and argue, following Broad[28], that this does not account for those sensations which are intrinsically pleasant (Sober and Wilson 1998, pp. 278-279). In their appraisal, "Butler's stone, besides being a fallacy, also has a false premise":

> It is false that there could be no pleasure unless the agent antecedently desired some external thing. And even if that connection between pleasure and desire were granted, it would not refute the hedonist's contention that people desire external things only because they think those things will satisfy their ultimate desire to gain pleasure and avoid pain. (p. 279)

In response to Sober and Wilson, it is important to note that Butler speaks of "all particular appetites and passions" as being towards external objects, and does not exclude the fact that self-love may pursue pleasure or a pleasurable state since, after all, its ultimate aim is the agent's happiness. Butler wishes, above all, to demonstrate that the plethora of human motives cannot be reduced to one, self-interest, and his theory of human nature as a system or constitution is an illustration of this. The premise of psychological hedonism which he attempts to refute is that all we ever seek is pleasure. In Butler's view, the fact that the pleasure deriving from our passions is our own does not entail that all we ever seek is pleasure nor that we are continually motivated by self-love:

> Every particular affection, even the love of our neighbour, is as really our own affection, as self-love; and the pleasure arising from its gratification is as much my own pleasure, as the pleasure self-love would have, from knowing I myself should be happy some time hence, would be my own pleasure. And if, because

[28] They refer to Broad 1930 at p. 66: "We must therefore distinguish between intrinsic pleasures and pains and the pleasures and pains of satisfied or frustrated impulse. [...] This kind of pleasure and pain is quite independent of the object of the impulse."

> every particular affection is a man's own, and the pleasure arising from its gratification his own pleasure, or pleasure to himself, such particular affection must be called self-love; according to this way of speaking, no creature whatever can possibly act but merely from self-love; and every action and every affection whatever is to be resolved up into this one principle. (Sermon XI, (7), p. 168)

Butler has three objections to the hedonist's theory. The first is that of self-destructive behavior or the reckless pursuit of pleasure. As previously mentioned, Butler considers a person who is prepared to ruin her life for the gratification of a particular passion as not having natural affection towards herself. She does not act in her best interest and, from the view of a cool outside observer, does not have enough reasonable self-love. The hedonist, however, could always respond that while she may fail in promoting her interest in Butler's sense of the word, her actions may still be motivated by what she deems to be her interest at the time.[29]

Butler's second objection concerns benevolent actions which we choose to do knowing they will go against our short and/or long-term interest. The hedonist's response to Butler is illustrated by Henson's example of Edna helping Ted: " 'Well, of course, it gives Edna pleasure to help him – at least it would have distressed her not to – and she does it *in order to gain that pleasure or avoid that distress*.' " (1988, p. 40). It could thus be argued that this is a case where we do what we want to do even though there is a conflict with other wants we have. We are again giving preponderance, or at least equal consideration through rational deliberation, to our own interest.

Butler's final objection claims that such a pleasure-seeking attitude would, in any event, be self-defeating: self-love in excess, wholly engrossing the self, fails to achieve its end, because it ignores the role of the passions in the

scheme of human nature and the prior suitableness of certain objects in satisfying those passions:

> *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. (Sermon XI, (9), p. 171)[30]

Sober and Wilson refer to this last objection as the "paradox of hedonism", which they consider a possible argument against ethical egoism, but not psychological egoism. "Even if this point entailed that people *should* not be hedonists, it would not follow that people are not hedonists *in fact*. Hedonism as a descriptive thesis needs to be distinguished from hedonism as a normative thesis. This so-called paradox has no bite with respect to the descriptive claim." (1998, p. 280)

Butler's core argument against psychological egoism and hedonism is that its representation of human nature is biased, and does not correspond to a realistic appraisal of our varied motivations:

> Therefore it is not a true representation of mankind to affirm, that they are wholly governed by self-love, the love of power and sensual appetites: since, as on the one hand they are often actuated by these, without any regard to right or wrong; so on the other it is manifest fact, that the same persons, the generality, are frequently influenced by friendship, compassion, gratitude; and even a general abhorrence of what is base, and liking of what is fair and just, takes its turn amongst the other motives of action. (our emphasis) (*Sermons*, The Preface, (21), p. 13)

---

[29] Duncan-Jones 1952, p. 62 suggests other possible reasons for the failing of self-love, such as lack of intelligence, education or knowledge of the world.

[30] See also Broad 1930 at p. 74: "Happiness which is deliberately sought generally turns out to be disappointing, and the self-conscious egoist divides his time between wanting what he has not and not wanting what he has."

As summarizes Joel Feinberg, "The view then that we are never after anything in our actions but our own pleasure–that all people are complete 'gourmets' of one sort or another–is not only morally cynical; it is also contrary to common sense and everyday experience." (1999, p. 500). When the egoist identifies the search for happiness (and the hedonist the pursuit of pleasure) as the ultimate motivator of all human action, he is in fact confusing the object of self-love with that of the particular passions.

Butler perhaps concedes to the egoist that when we consider the effects of our actions on ourselves and on others, we give equal or more weight at all times to our interests *versus* those of others. In this sense, what Johnson calls "purely altruistic" actions, with no concern at all for one's self-interest, can never occur. However, it is questionable whether this is an uncontested fact about human nature. What about the rescuer who risks his life and the doctor treating lepers? Are they really acting with their own interest in view? We could just as well question whether "purely selfish" actions ever occur and whether there isn't always a degree of consideration towards others to be found in our actions. According to empirical studies quoted by Sober and Wilson, in situations where there is no conflict between our welfare and that of others, we choose the action which not only benefits ourselves, but which avoids harming or benefits others as well: "Those who think that human behavior makes the egoism hypothesis obvious should think again." (1998, pp. 247-248).

Duncan-Jones thinks that Butler's eighteenth century optimistic view of human nature prevents him from concluding that there are individuals so constituted that they can only gain satisfaction through power, exploitation or contention and the suffering of others (1952, p. 113). We reiterate in response to this comment that, without being naïve, Butler believes that the plurality of human motives cannot be reduced to a single motive. In his view, we are alternately motivated by the particular passions, the rational principles of self-love and benevolence, and the moral considerations of our conscience.

Furthermore, Butler would respond that such a person could not possibly obtain true happiness in this manner. As Henson claims:

> Butler does not have "proofs" that we are not always selfish, nor does he suppose that the question is one which should be capable of being so settled. (See his first long footnote to Sermon I.) The one thing that I think he has proven – that self-love cannot be our only motive – opens us to the *possibility* that any of an indefinite range of motives may work in us; and he reminds us both how regardless of our own welfare they may be and how varied may be those that serve it. (1988, p. 57)

Moreover, it is true that for Butler, duty and interest coincide; in Johnson's words, "the path of true morality is also the path of our own self-interest", and "my own flourishing *depends* on the flourishing of other human beings" (1992, p. 249). Hence, we could conclude along with the egoist, all the while concurring with Butler's theory of human nature, that no deliberate actions are entirely devoid of self-interest[31], and that self-love always occupies as important a place as any other motivator of human action. This is what the notorious "cool hour passage" which will be discussed in the following section seems to endorse:

> Now all this confusion might easily be avoided, by stating to ourselves wherein the idea of self-love in general consists, as distinguished from all particular movements towards particular external objects; the appetites of sense, resentment, compassion, curiosity, ambition and the rest. When this is done, if the words *selfish* and *interested* cannot be parted with, but must be applied to everything; yet, to avoid such total confusion of all language, let the distinction be made by epithets: and the first may be called cool or settled selfishness,

---

[31] See Scott-Taggart 1968 at p. 28: "There are, then, actions done out of disinterested motives, but the question is whether there are any actions which are done out of *purely* disinterested motives." We reiterate that Butler would perhaps respond: "But are there any purely *interested* motives for such sentient and social beings as us?"

> and the other passionate or sensual selfishness. (*Sermons*, The Preface, (35), p. 21)

Let us now study the debate over the controversial "cool hour passage" in Butler's *Sermons*.

2.2 The "Cool Hour Passage"

There is a controversial passage in the *Sermons* which has lead certain commentators to interpret Butler as an ethical egoist. It is known by the initiated as the "cool hour passage":

> Let it be allowed, though virtue or moral rectitude does indeed consist in affection to and pursuit of what is right and good, as such; yet, that when we sit down in a cool hour, we can neither justify to ourselves this or any other pursuit, till we are convinced that it will be for our happiness, or at least not contrary to it. (Sermon XI, (20), p. 182)

In this passage, Butler seems to be saying that no action can be rationally justified unless it is in our interest, or at least not against our interest. It is unclear whether this is a statement about human nature and moral psychology, or whether Butler is advocating ethical egoism, *i.e.* self-interest as the ultimate justification of moral action. Is he conceding to Hobbes that the human makeup is inherently selfish or at least self-centered, or is he saying that we should not act contrary to our happiness?

Butler prefaces the "cool hour passage" with the following:

> [...] and there can no access be had to the understanding, but by convincing men, that the course of life we would persuade them to is not contrary to their interest. It may be allowed, without any prejudice to the cause of virtue and religion, that our ideas of happiness and misery are of all our ideas the nearest

> and most important to us; that they will, nay, if you please, that they ought to prevail over those of order, and beauty, and harmony, and proportion, if there should ever be, as it is impossible there ever should be, any inconsistence between them: though these last too, as expressing the fitness of actions, are real as truth itself. (our emphasis) (Sermon XI, (20), p. 181-2)

It appears from the above that Butler does not believe we can be convinced to act solely in the name of duty, unless we are assured that such course of action is not contrary to our interest. However, in the same breath, he states that it is impossible for there to be a contradiction between the ideas of happiness and misery and those expressing the fitness of actions, *i.e.* order, beauty, harmony and proportion. We have seen (and will explore further in the next chapter) that Butler's notion of due proportion or harmony involves an equilibrium between the different principles of human nature, each fulfilling the role for which it was designed. Although this is an ethical (and esthetical) ideal, similar to the attainment of truth, Butler realizes that in this world, we suffer from ignorance, imperfection and striving (See Sermon III, (2), footnote 2, p. 63). Ultimately, Butler does not believe that a conflict between the paths which conscience and self-love lead to respectively, *i.e.* virtue and interest, is possible:

> It is manifest that, in the common course of life, there is seldom any inconsistency between our duty and what is *called* interest: it is much seldomer that there is an inconsistency between duty and what is really our present interest; meaning by interest, happiness and satisfaction. Self-love, then, though confined to the interest of the present world, does in general perfectly coincide with virtue; and leads us to one and the same course of life. (Sermon III, (8), p. 67)

Butler pursues:

> Reasonable self-love and conscience are the chief or superior principles in the nature of man: because an

> action may be suitable to this nature, though all other principles be violated; but becomes unsuitable, if either of those are. Conscience and self-love, if we understand our true happiness, always lead us the same way. (Sermon III, (9), p. 68)

This last passage lead Thomas H. McPherson to call Butler a "rational egoist" or "ethical eudaemonist" in the *Sermons*, but an "empirical intuitionist" (or deontologist) and a "transcendental utilitarian" in the *Analogy*, which comprises the *Dissertation* (1948, Part I, p. 330; 1949, Part II, pp. 11, 16-17). According to McPherson, Butler deems the production of happiness for the agent to be the ground or criterion of rightness of our acts (1948, Part I, p. 327). In our view, although Butler does consider the pursuit of happiness as an obligation which is the nearest and dearest to us, he does not consider it to be the basis of morality. Butler believes that there is a universal standard of morality (which includes justice, veracity, regard to common good and prudence), but that "good is indefinable", to quote George Edward Moore (1971, p. 8), disciple of Henry Sidgwick, who was in turn a disciple of James Martineau, Butler's disciple. As McPherson himself says, speaking of both the *Sermons* and the *Analogy*:

> Right and good are not to be defined in terms of anything else. [...] Butler would probably have been very glad to echo Professor Moore's well-known remark [...]: "If I am asked 'What is Right?' my answer is that right is right, and that is the end of the matter. Or if I am asked, 'How is right to be defined?' my answer is that it cannot be defined, and that is all I have to say about it.". (1949, Part II, pp. 18-19)

Right action is revealed to us on a case-by-case basis through the voice of our conscience, which is the guide of life, and our witness of the law within us.

Henry Sidgwick, in *The Methods of Ethics*, reads Butler's "cool hour passage" as granting dual authority in human nature to conscience and

reasonable self-love. Sidgwick reasons, firstly, that since interest or one's happiness is a manifest obligation to Butler, it is a moral "ought", a "categorical imperative", and part of our duty:

> The word 'ought' thus used is no longer relative: happiness now appears as an ultimate end, the pursuit of which – at least within the limits imposed by other duties – appears to be prescribed by reason 'categorically', as Kant would say, *i.e.* without any tacit assumption of a still ulterior end. And it has been widely held by even orthodox moralists that all morality rests ultimately on the basis of "reasonable self-love"; *i.e.* that its rules are ultimately binding on any individual only so far as it is his interest on the whole to observe them. (our emphasis) (1981, p. 7)

Sidgwick believes that, for Butler, self-love is "one of two superior and naturally authoritative impulses, the other being Conscience", and that in the "cool hour passage", Butler "even concedes that it would be reasonable for Conscience to yield to it, if the two could possibly conflict." (1981, p. 366). Sidgwick concludes that there is a duality of the regulative principles in human nature, *i.e.* "Universalistic Hedonism" and "Egoistic Hedonism". Since he puts Butler's conscience and self-love on an equal footing, Sidgwick then deduces a dualism of the practical reason, reason being divided against itself by two ultimate and fundamentally contradictory principles, virtue and self-interest. To Sidgwick, this is Butler's most important contribution to the advancement of theoretical ethics. Sidgwick further states that, according to Butler, self-love is not subordinate to conscience in theory, but in practice, the dictates of conscience being the more clear, certain and probable, whereas we often cannot determine what will lead to our happiness:

> And it is by appealing to the superior certainty with which the dictates of Conscience or the Moral Faculty are issued, that Butler maintains the practical supremacy of Conscience over Self-love, in spite of his admission (in the passage before quoted) of theoretical

> priority in the claims of the latter. A man knows certainly, he says, what he ought to do: but he does not certainly know what will lead to his happiness. (1981, p. 200)[32]

However, it should noted that when Butler describes the conflict between the "obligation from supposed interest" and the "obligation to virtue" (he speaks of "supposed" interest since he does not believe that there could be a conflict between virtue and our true interest), he concludes that only the obligation to virtue remains. Conscience's authority, the key element Butler considers to be missing from Shaftesbury's moral theory in his *Inquiry concerning Virtue*[33], prevents us from being "under two contrary obligations". It is not the commands of conscience that are more certain and known, as deems Sidgwick, but the natural authority of the principle which obligates and binds:

> But the obligation on the side of interest really does not remain. For the natural authority of the principle of reflection is an obligation the most near and intimate, the most certain and known: whereas the contrary obligation can at the utmost appear no more than probable; since no man can be *certain* in any circumstances that vice is his interest in the present world, much less can he be certain against another: and thus the certain obligation would entirely supersede and destroy the uncertain one; which yet would have been of real force without the former. (our emphasis) (*Sermons*, The Preface, (26), pp. 15-16)[34]

---

[32] See also Broad 1930 at p. 80: "On the other hand, the dictates of conscience are often quite clear. Thus we can be far more certain about what is right than what is to our own ultimate interest; and therefore, in an apparent conflict between the two, conscience should be followed since we cannot be sure that this is not really to our own interest."

[33] See the *Sermons*, The Preface, (26), p. 15: "The not taking into consideration this authority, which is implied in the idea of reflex approbation or disapprobation, seems a material deficiency or omission in *Lord Shaftesbury's Inquiry concerning Virtue*."

[34] See A.E. Taylor 1926 at p. 299: "The contrast [Butler] makes is simply the contrast between the *uncertainty* that vice will be to our interest and the absolute *certainty* that the moral law carries its authority with it. The point is simply that this authority is 'certain and known'."; and Shaver 1999 at p. 120: "Butler does not seem to make the supremacy of

John Kleinig is of the opinion that Butler's "cool hour passage" is a descriptive statement regarding motivation, and does not address the issue of the foundation of morality. In other words, our nature is so constituted that, even after serious consideration of a moral dilemma, we will only do our duty if we feel that it will have a positive (or in the very least a non-negative) impact on our happiness. Kleinig's thesis helps explain why Butler argues so strongly in favour of the coincidence of duty and interest, in order to convince his congregation that their duty is in their long-term interest. However, it does not imply that Butler is indicating that self-love is, nor that it ought to be, the standard of right conduct. That is of the office of conscience alone:

> Butler is not saying that an action is right only if it can be seen to be conducive (or at least not contrary) to our happiness. He has a quite different point in mind, namely, that if we carefully reflect on any virtuous action, we will be able to justify our actual performance of it only if we are convinced that this will not be detrimental to our happiness. Butler thus distinguishes the question of performing an action from the question of whether it is right or wrong.
>
> It is one thing to decide what would be the virtuous thing to do; it is another thing to decide to do the virtuous thing. It is with the latter that the "cool hour" passage is concerned. (Kleinig 1979, p. 406)[35]

Contrary to Sidgwick and along the same line of thought as Kleinig, Alan R. White regards the edicts of Butler's conscience as subordinate to those of self-love in practice, in that they have less motivational strength, but not in theory, for as he states, "Butler never doubted that conscience was the only

---

conscience conditional on our greater certainty of its pronouncements. It is, instead, conditional on the 'claim' of conscience to be superior".

[35] See also A.R. White 1952, at p. 337: "This does *not* mean that the ground of rightness is conduciveness to happiness, it does not even mean that the motive of self-love is a good motive–the good motive is "affection to and pursuit of what is right and good as such"–it simply means that as a motive to action, self-love is more influential than the dictates of conscience."

arbiter of virtue" (1952, p. 338). Unlike McPherson, White does not think that Butler is tackling the question of rational justification of ethical action in the "cool hour passage", only that of motivation:

> The word "justify" in this passage is not, in my opinion, being used in an ethical sense, but means that happiness-loving men can see no reason to do action A, that is, cannot "justify" to themselves their choice of action A, unless it plainly appears to be productive of their happiness. The fact that the action is right is not for them sufficient justification or reason to pursue it. (1952, p. 337)

In any event, since Butler believed that duty and interest coincide, in his mind there was little harm done in granting the motivational force of self-love to his congregation, for at the time he was preaching his *Sermons* it was fashionable to believe that human nature was inherently selfish. Indeed, the "Broad-Taylor" view of the "cool hour passage" is that Butler "is not here asserting his own view, but is simply making a hypothetical concession to an imaginary opponent" (Broad 1930, p. 80):

> The phraseology here, "it may be allowed", "if you please", "if there ever should be, as it is impossible there ever should be", "let it be allowed", makes it, to my mind, almost certain, that the position is meant to be one which Butler himself does not accept, but is content to assume for the purpose of reasoning with an audience who regard "self-love" as the one rational rule of conduct. He had said at the beginning of the discourse, "there shall be all possible concessions made to the favourite passion, which hath so much allowed to it, and whose cause is so universally pleaded; it shall be treated with the utmost tenderness, and concern for its interests". We must therefore expect a touch of "irony" in the sermon. It is monstrous to discuss the incriminated passage without taking into account Butler's sarcastic words about the indulgence he proposes to show the "favourite passion" of his auditors. (A. E. Taylor 1926, p. 295, footnote 1)

However, if the "cool hour passage" was a concession on Butler's part, he was by no means agreeing with a reductionist view of human nature whereby all human action is motivated by selfishness.

There is another interpretation of the "cool hour passage" with respect to the term "justification" for action which distinguishes rationality and morality, or non-moral *versus* moral ultimates. Here is how Edmund Leites phrases it: "The rational requirement that we do the bidding of self-interest exists independently of the moral requirement that we obey conscience. [...] Unlike Kant, Butler does not identify the supreme moral requirement upon action with its supreme rational requirement." (1975, p. 45). This view opposes that of Sidgwick, who grounds morality in rationality, *i.e.* in what is considered to be ultimately reasonable. Furthermore, because Sidgwick considers the production of happiness to be ultimately reasonable, it becomes the criterion of rightness. Butler's conscience, as the supreme authoritative moral faculty, purveys our entire nature before instructing us on which course of action is appropriate. Self-love, as a principle of reflection concerning specifically the agent's happiness, uses instrumental rationality to achieve its end. However, in cases where conscience and self-love may conflict, *i.e.* where the moral course of action differs from that of rational self-interest (which, for Butler, is impossible, so much of this discussion is academic for him); hypothetically, we would have to choose between morality and happiness. What we have here is not a theoretical contradiction at the root of reflection itself, but rather a problem of choice.

This brings us to a final interpretation of the "cool hour passage". According to William J. Norton and Hastings Rashdall (a disciple of Sidgwick), for Butler, virtue is an essential element of real *versus* "supposed" self-love, and virtuous behaviour part of our true happiness. Conscience is a means of completion of the actual self (Norton 1940, p. 97 *ss.*), and goodness of the essence of the highest happiness (Rashdall 1971, p. 62). There is no inherent

contradiction between Universalistic and Egoistic Hedonism as deems Sidgwick, *i.e.* between the fact that it is right and reasonable to promote the common good as well as our own good; although the two may, in certain cases, lead to different courses of action and call for decision-making on our part. Since it is of the nature of the good that it is of intrinsic worth to us, we must consider it as a good not merely to others, but to ourselves as well, and thus, argues Rashdall, conducive to our own happiness. Butler would probably concur with this view:

> For does not all this kind of talk go upon supposition, that our happiness in this world consists in somewhat quite distinct from regards to others; and that it is the privilege of vice to be without restraint or confinement? Whereas on the contrary, the enjoyments, in a manner all the common enjoyments of life, even the pleasures of vice, depend upon these regards of one kind or another to our fellow-creatures. (Sermon III, (7), p. 65)

In the words of Ramon M. Lemos, most of us are averse to thinking of ourselves as being morally bad, which implies a belief that our "complete good consists not only of a nonmoral component consisting of [...] being happy but also of a moral component consisting of [...] being morally good" (2003, p. 212).

In conclusion, reasonable self-love and conscience are the "chief or superior principles" in human nature because we cannot violate them without violating our entire nature. The various interpretations of the "cool hour passage" suggest that, while Butler believes virtue and interest coincide in the end, he considers it a fair question to ask "Why should we be moral?" when we find that the path of virtue does not appear to be in our interest. However, Butler's response to this query is that, despite our special obligation of care towards ourselves which spurs us to "justify" our acts in the sense of self-interest, we have a manifest obligation to follow our conscience which supersedes this prior obligation since conscience alone claims authority.

Finally, in response to T.H. McPherson who claims that Butler developed from an "ethical egoistic eudaemonist" in the *Sermons* to an intuitionist (deontologist) in the *Analogy*, it is true that in the *Dissertation* Butler speaks of the virtue of *prudence* rather than of self-love. However, they both seem to refer to one and the same thing, *i.e.* "a due concern about our own interest or happiness" (*Dissertation*, (6), p. 251), which is approved by conscience as a moral obligation or duty. As observes Darwall, "In the *Dissertation* Butler distinguishes between self-love and the judgment that a person ought to promote his interest"[36] (1995, footnote 31, p. 260):

> First-order conscience approves or disapproves not only of acts but also of *principles* (or principled conduct). [...] And Butler holds that self-love is one of the principles we reflectively approve of, at least when it is properly tempered: [...] Butler does say that what we are approving of here is "conduct", but he appears to mean conduct motivated by a certain principle – specifically, by "a due concern about our own interest". A second example of a principle approved by conscience is, of course, benevolence. (1995, p. 273)

In our opinion, the "cool hour passage" is of too great significance and consequence to Butler's hierarchical moral theory for it to be dismissed as a concession for the sake of argument. It is precisely for this reason that it has been the subject of such debate. Even though Butler's thought may, quite naturally, have evolved to some extent over the seven-year period between the publication of the second edition of the *Sermons* and of the *Analogy*, we believe that it is possible to interpret them as one system of ethics and maintain consistency. If we study Butler's text closely, we note that he uses the terms "cool hour", "justify", "convinced", "access to the understanding" of,

---

[36] "This approbation and disapprobation are altogether different from mere desire of our own, or of their happiness, and from sorrow upon missing it. For the object or occasion of this last kind of perception is satisfaction or uneasiness: whereas the object of the first is active behaviour. In one case, what our thoughts fix upon is our condition: in the other, our conduct." (*Dissertation*, (6), p. 252)

"persuading" men along with the "ideas of": "interest, happiness and satisfaction" and self-love; and opposes the above terms with "virtue or moral rectitude", "affection to and pursuit of what is right and good, as such", "cause of virtue and religion", "order, and beauty, and harmony, and proportion", "fitness of actions" and duty.

If we situate the "cool hour passage" within the context of Butler's eleventh sermon, *Upon the Love of our Neighbour*, from which it is taken, we note that the entire sermon aims to demonstrate that there is no particular contrariety between the principles of self-love and benevolence. Pursuing one's happiness (or good) is the object of self-love; whereas pursuing the happiness of others (or the common good) is the object of benevolence. Not only are the two principles compatible according to Butler, but they are rational teleological principles and in no way constitutive of morality. Pursuing the happiness of others concerns benevolence, not conscience. In addition, self-love and benevolence are principles which conscience itself approves of and which pose no threat to its dictates.

Butler was well aware of the debate between duty and interest and, in our view, not only was he not ignoring it, but he was attempting to address the issue directly in his eleventh sermon. Butler appears to be confronting self-love/rational thought, with virtue and morality/conduct ("pursuit", "cause of life", "actions"). Butler situates self-love on the level of a purely rational principle, whereas conscience is more intuitive, providing access to universal standards of morality, the fitness of actions, the ideals of order, beauty, harmony and proportion, not through the understanding, but through an inner voice. This guide of life reveals a realm which is, however, as "real as truth itself" (Sermon XI, (20), p. 182): "Acting, conduct, behaviour, abstracted from all regard to what is in fact and event the consequence of it, is itself the natural object of the moral discernment; as speculative truth and falsehood is of speculative reason." (*Dissertation*, (2), p. 248).

To conclude, it is doubtful that, for Butler, morality could ultimately be justified purely in rational terms. He argues on the level of reason and is keenly conscious of the need to "convince" his congregation, all the while carrying the torch of his faith which comforts him with the knowledge that, in the end, the paths of virtue and interest coincide.

In the next chapter, we will study more closely Butler's theory of moral conscience, along with the different arguments he brings forward to illustrate its authority.

CHAPTER 3

Butler's Theory of Moral Conscience

We saw in our first chapter that Butler views human nature as a system or constitution, with each principle playing a particular role in the conduct of an individual. Butler states in Sermon II, *Upon Human Nature*, that to act in conformity with human nature as a whole involves recognizing the supremacy of our conscience. According to Butler, conscience is supreme since, by natural right, it "ought to rule". Conscience's authority places it at the top of the hierarchy of principles in human nature. In this chapter, we will examine the nature and the source of conscience's authority for Butler.

Conscience presides over the other principles in our nature and guides our conduct. If we are uncertain how to act in a particular situation, such as in a case of conflict between self-love and benevolence, conscience will not necessarily provide the answer. However, if one of the courses of action we are contemplating is "wrong" in that it contravenes our duty (or "right" in that it corresponds to a duty), conscience will speak out loud and clear and let us know which course of action is unacceptable (or imperative), unless we are willing to suffer the consequences of knowingly violating our conscience. Conscience acts as a morality check on our motives, intentions and actions.

3.1 The Role of Conscience in Human Nature

Butler's theory of human nature as a system or constitution bestows upon conscience authority over the other principles. Contrary to brutes, who act primarily from the passions and instinctually for self-preservation and the survival of the species, we possess the capacity to reflect on our actions and motives, as well as those of others, then pass judgment on them and on the

agent, and act accordingly. It is because we have this faculty of moral approval and disapproval that we are capable of being under moral government:

> That which renders beings capable of moral government, is their having a moral nature, and moral faculties of perception and of action. Brute creatures are impressed and actuated by various instincts and propensions: so also are we. But additional to this, we have a capacity of reflecting upon actions and characters, and making them an object to our thought: and on doing this, we naturally and unavoidably approve some actions, under the peculiar view of their being virtuous and of good desert; and disapprove others, as vicious and of ill desert. That we have this moral approving and disapproving faculty, is certain from our experiencing it in ourselves, and recognizing it in each other. (*Dissertation*, (1), pp. 246-247)

According to Butler, it is a matter of experience that conscience exists. We have proof of its existence through our own judgment of good and ill desert and of moral character, as well as our condemnation of vice and commendation of virtue. For example, we distinguish between mere harm and injury deserving just punishment. Our system of rewards and punishments enforces the commands of our conscience and, furthermore, self-condemnation reveals the voice of our conscience when we have acted with malicious intent.

Butler's conscience is a rational faculty through which we have access to a "universally acknowledged standard" of virtue, *i.e.* "justice, veracity and regard to common good" (*Dissertation*, (1), p. 248) and "prudence" ((6)-(7), pp. 251-253). In Sermon XII, Butler states that "there are certain dispositions of mind, and certain actions, which are in themselves approved or disapproved by mankind, abstracted from the consideration of their tendency to the happiness or misery of the world; approved or disapproved by reflection, by that principle within, which is the guide of life, the judge of right and wrong." The examples Butler gives of "particular obligations" besides "doing good, or producing happiness" are greatness of mind, fidelity, honour and strict justice

as opposed to treachery, indecencies and meanness of mind (Sermon XII, (31), footnote 1, pp. 199-200). Moreover, the rational principles of self-love and benevolence are subject to the scrutiny and final approval of conscience.

Butler's conscience is a rational faculty with intrinsic moral authority, in that its dictates carry with them the effect of a moral obligation. We may ignore our conscience and decide not to do our duty, but we know intuitively that it is our duty. However, as previously discussed, conscience does not possess strength as it possesses authority. The passions or supposed self-love may in many cases be stronger than the temptation to the path of virtue. However, conscience "ought" to rule in human nature if we are to accomplish the purpose for which it was implanted in us, *i.e.* moral agency. According to Butler, had there been no difference between the principles in human nature except that of strength, there would have been no rational foundation of moral judgment.

Hence, conscience is the faculty which makes it possible for us to choose to do good and restrain from mischief. Due to the possession of conscience, man is a "law unto himself". In Butler's own terms, rationality here includes "both the discernment of what is right, and a disposition to regulate ourselves by it" (Sermon XII, (9), p. 189):

> For when the Gentiles, which have not the law, do by nature the things contained in the law, these, having not the law, are a law unto themselves;
>
> Which shew the work of the law written in their hearts, their conscience also bearing witness, and their thoughts the mean while accusing or else excusing one another; (Rom. ii.14-15)[37]

[37] The New Testament, *A Reader's Guide to the Holy Bible*, King James Version, 1972, Thomas Nelson Inc., Nashville, Tennessee, p. 141.

a) Conscience as Authoritative

Let us consider the above-quoted passage, in order to understand the import of the authority of conscience in Butler's theory. In Sermons II and III, Butler uses this quotation from *To the Romans* to argue against the Deists that, whether we believe in revelation[38] or not, we cannot escape the authority of conscience.

Conscience is, in Butler's words, "the guide of life", the voice of God within us. Let us take the Christian commandment "Thou shalt not kill" thy fellow man (Rom. xiii.9). Disregarding for now our religion or system of punishment, and barring exceptional circumstances such as self-defense or the protection of our kin, it is inconceivable to us that we should violate this law without suffering not only condemnation by others, but self-condemnation as well. This dictate is one of the universally accepted moral principles. Thus, even though we may not "have the law" through divine revelation, its commandments are all the same forged in our hearts. When, by our nature, we do not kill, and thus respect the law, we are a law unto ourselves. We are acting morally by following a moral standard revealed to us through our conscience. When we violate this law, our conscience haunts us with the knowledge of our act, even though it may be our secret. Not only is moral conscience a "moral approving and disapproving faculty" as concerns actions and characters, but it is a faculty which approves of itself upon reflection. This is its authority:

> [...] upon a survey of actions, whether before or after they are done, it determines them to be good or evil; and also because it determines itself to be the guide of action and of life, in contradistinction from all other faculties, or natural principles of action, in the very same manner as speculative reason *directly* and

[38] Butler was well aware that, as a rule, his congregation believed in God and an afterlife. It was Christian revelation which was contested at the time by the Deists.

> naturally judges of speculative truth and falsehood; and at the same time is attended with a consciousness upon *reflection*, that the natural right to judge of them belongs to it. (*Dissertation*, (1), footnote 1, p. 247)

No matter how strong the passions may be, and even once they have urged us to act and we have done so, their power cannot compare with the authority of conscience. Their influence is localized, concentrated and usually short-lived. Conscience, on the other hand, may lie dormant in the ordinary course of events: if morality is not in question, we may be entirely unaware of its existence. However, when we are faced with a moral dilemma, the authority of conscience immediately reveals itself to us. If we choose all the same to act immorally, though we may escape external sanction, and according to Butler, regardless of our own feelings concerning our decision, we will recognize the reasons of conscience as authoritative. This is the distinction between *mere power* and *authority*: "between, that is, a principle's having strength and its being one by which an agent *should* govern himself" (Darwall 1995, p. 255). Just as we may ask ourselves "Why should I be moral?" before deciding to commit an immoral act and put aside the moral law, we will be conscious of the immorality of the act once it is done and of the authority of the law. At that point, when we see ourselves as having acted immorally, unless we are depraved, our idea of self will be diminished, since the act we previously identified with while pushing morality aside is now ours, appropriated by us through our doing. In addition, according to Butler, the actual consequences of our actions are irrelevant (eg. failure to succeed in our enterprise); it is the intended consequences which determine the morality or immorality of an act. Moral conscience is our shadow which follows us wherever we go, whatever we do: "Had it strength, as it has right: had it power, as it has manifest authority; it would absolutely govern the world." (Sermon II, (14), p. 57). As A. E. Taylor writes:

> [Butler's] point is simply that we know with certainty that we have specific moral obligations, and that uncertainty whether neglect of them will affect our "interest" does nothing to destroy this certainty of their reality. Their reality would, in fact, remain, even if I could be certain, as I cannot be, that vice will do nothing to diminish my happiness, or even that a vicious act will augment it. If you grant the intrinsic "authoritativeness" of the moral law, even certainty that I should gain by violating it would not in any way affect the other certainty that I ought not to violate it, any more than certainty that I shall be rewarded by the Government for committing a convenient crime would make the commission of the crime a lawful act. (1926, p. 300)[39]

b) Conscience as a Moral Faculty

How is conscience a moral faculty, in Butler's view? Let us examine the passages in the *Sermons* where he defines conscience and its function.

In the Preface to the *Sermons* and subsequently, Butler uses "conscience" and "reflection" interchangeably. As we saw, conscience is the faculty which renders us, in contradistinction to brutes and machines, capable of moral agency: "A machine is inanimate and passive: but we are agents. Our constitution is put in our own power. We are charged with it; and therefore are accountable for any disorder or violation of it." (*Sermons*, The Preface, (14), p. 11)

But what is our "constitution" which is "within our power" in Butler's eyes? Butler has borrowed these notions, which are key to his understanding of conscience, from the Stoics, notably from Epictetus, who inspired Butler to

---

[39] See also Duncan-Jones 1952 at pages 87-89: "The assumption that the presence of a moral quality is always in itself a reason for some course of action may be named 'the principle of the intrinsic stringency of moral qualities'. This is the principle that their 'stringency' is 'prior to all will whatever', and does not arise from their attractiveness or unattractiveness to anyone. [...] the principle of intrinsic stringency seems to be a more precise and explicit

describe conscience as a "moral approving and disapproving faculty" (*Dissertation*, (1), pp. 246-247). In the *Enchiridion*, Epictetus declares that our opinions, movements towards a thing, desires, aversions, in short, whatever are our own acts, are "in our power". The things "not in our power" are our body, property, reputation, offices, in short, whatever does not stem directly from our own acts (trans. George Long 1991, I, p. 11). When Butler states that we are agents since our constitution has been placed in our power, he is further imparting the statement with moral content. In addition to having passions and aversions, we have the capacity to choose to move towards an object(ive) or not and to think in a certain way; in essence, we are free agents who are responsible for our acts.

According to Epictetus, only one faculty contemplates and consequently approves or disapproves of itself: it is the reasoning faculty. It is this faculty alone which judges the right use of the appearance of things, *i.e.* which exerts the powers of pursuit and avoidance, of desire and aversion. The gods placed in us a capacity to make the best use of what is in our power, and Epictetus gives the examples of deciding not to betray a secret (*Discourses: Book I*, trans. Robert F. Dobbin 1998, Chapter I: "Concerning What Is In Our Power and What Is Not"). These examples evidently illustrate a faculty of moral, and not just rational, decision-making. What is interesting here for Butler is the following: there is an approving/disapproving faculty which, having been put in our power, renders us capable of moral agency, and furthermore, of subsequent reflection on this moral agency. Thus, conscience is characterised by an unreflective level, *i.e.* a spontaneous moral judging more concomitant with moral sentiment; as well as a reflective level, *i.e.* moral reasoning (which we sometimes misuse, arguing away our duty). The unreflective level of conscience has the following advantage over the reflective level: it provides us

---

statement of what Butler had in mind when he spoke of 'authority'. [...] The authority of conscience lies in the intrinsic stringency of the moral quality".

with immediate access to our duty, without the tricky detour of reasoning, which often, "reasons duty away". However, it is through the reflective level of conscience that we are conscious of its authority, and of the validity of its pronouncements.

Following Lord Shaftesbury, Butler also qualifies conscience as "reflex approbation or disapprobation" (*Sermons*, The Preface, (26), p. 15). Contrary to conscience as reflection, which describes conscience as a reasoning faculty capable of directly motivating moral action, reflex approbation/disapprobation has an immediacy, inevitableness, or intuitive quality. We commend or condemn our actions and motives as well as those of others spontaneously, by a gut reaction or feeling. What Butler means by reflex approbation / disapprobation is revealed in the following passage – conscience exercises itself when morality is in question, without even being consulted:

> But there is a superior principle of reflection or conscience in every man, which distinguishes between the internal principles of his heart, as well as his external actions: which passes judgment upon himself and them; pronounces determinately some actions to be in themselves just, right, good; others to be in themselves evil, wrong, unjust: which, without being consulted, without being advised with, magisterially exerts itself, and approves or condemns him, the doer of them, accordingly: and which, if not forcibly stopped, naturally and always of course goes on to anticipate a higher and more effectual sentence, which shall hereafter second and affirm its own. [...] It is by this faculty, natural to man, that he is a moral agent, that he is a law to himself: but this faculty, I say, not to be considered merely as a principle in his heart, which is to have some influence as well as others; but considered as a faculty in kind and in nature supreme over all others, and which bears its own authority of being so. (our emphasis) (Sermon II, (8), pp. 53-54)

This is the most complete description Butler gives of conscience: conscience is a superior principle of reflection, as well as the supreme faculty

which renders us moral agents, by immediately exerting its approval or disapproval of intentions and actions, while at the same time revealing to us its authority over all the other principles in human nature.

Butler's description of conscience raises certain questions. Firstly, how is conscience a principle of reflection? Unlike the particular passions, but similarly to self-love and benevolence, conscience does not rest in external objects as its end, but is a regulating principle, surveying the passions and lower principles of self-love and benevolence and passing judgment on the various courses of action available. Its end is moral action, and its ideal goal (though only partially attainable for such imperfect creatures as us humans) is virtue, just as the end of self-love is acting in our interest in order, ultimately, to attain happiness, and that of benevolence is goodwill to our fellow creatures in view of the common good. While benevolence and self-love are similarly superior principles of reflection, however, they are not "faculties" and thus do not govern our entire constitution. Conscience as a "faculty was placed within to be our proper governor; to direct and regulate all under principles, passions and motives of action." (Sermon II, (15), p. 57).

Secondly, what exactly is this "approving or disapproving" faculty? Butler's theory is certainly not an emotivist one, with conscience as a purely psychological disposition or mode of feeling; nor is his idea of conscience that of an affection towards others for their welfare (which, rather, would be benevolence); nor is it the equivalent of Hume's sympathy. Butler's conscience is, above all, a moral faculty which makes practically effective judgments. Here is how Darwall explains Butler's faculty of "reflective approval and disapproval":

> Like Hutcheson, Butler grounded moral obligation in a "reflex approbation" – what Shaftesbury and Hutcheson had dubbed "moral sense" – but with a crucial difference. Whereas Hutcheson claimed that moral

> sense can have no direct practical role, Butler insisted that the faculty of reflective approval and disapproval – "conscience", or the "principle of reflection", as he called it – has an intrinsic *practical authority*. (1995, p. 245)

Conscience judges particular cases directly and intuitively, without recourse to general principles of morality, and at the same time, it can apply these general principles of morality by reasoning to particular cases directly. In Leites' terms:

> [Butler] thinks that conscience ordinarily judges particular cases *directly*, without recourse to general principles which it then has to apply by reasoning to particular cases. This direct judgment is usually adequate: [...] Learned men may seek to work out the general principles which explain the judgments of any honest man's conscience; but the honest man himself does not need to reflect on such principles and their application. He can judge the case at hand directly. (Leites 1974, p. 54)

Butler attempts to avoids the debate which became popular in the mid-eighteenth century and is still very much a polemic today, disputing whether moral sentiment or moral reason is the foundation of morality. Butler is of the opinion that the moral faculty embraces both moral sentiment and moral reason: "whether called conscience, moral reason, moral sense, or divine reason; whether considered as a sentiment of the understanding, or as a perception of the heart; or, which seems the truth, as including both."[40] (*Dissertation*, (1), p. 247). As resumes W. R. Matthews in his introduction to Butler's *Sermons*:

---

[40] This ressembles Pascal's paradox: "*Le coeur a ses raisons que la raison ne connaît point.*" According to Schneewind 1998, Butler had read Pascal (see footnote 25 at p. 343). Also see Duncan-Jones 1952, at pp. 75-76: "But Butler's deliberately elusive language does not authorise us to say that an act of conscience is primarily cognitive – a 'perception' or act of 'understanding' – rather than affective or conative – a 'sentiment' or a movement of the 'heart'."

> *The nature of Conscience.* Butler's position on this question may be regarded as a kind of *via media* between Rational and Aesthetic Intuitionism. He seems to agree with Shaftesbury that there is a special moral faculty, but for him it is more than feeling or instinct. It is a "principle of reflection", it partakes of the nature of reason. Nevertheless the Conscience is not the same as the theoretical or pure reason, as the Rational Intuitionists seemed to imply. (1967, p. xxiii)

According to James Seth, Butler's conscience is a purely rational principle which lacks the aesthetic and emotional element found in the moral sense of Shaftesbury and Hutcheson (1912, p. 201). However, in Albert Lefevre's view, Butler consciously avoids following in the footsteps of the rational intuitionists through his choice of method:

> Now, Butler, following, on the whole, the lead of Hobbes's opponents, seeks to demonstrate that morality is grounded in the peculiar nature and constitution of man. He sees that there are two methods which may be employed, one which aims at showing that morality is part of the "nature of things", the other that it is grounded in the peculiar constitution of human nature. Cudworth and Clarke had adopted the former, but Butler properly proceeds according to the latter method. (A. Lefevre 1899, p. 130)

In our opinion, whether we take Butler's theory of conscience to express a paradox, a middle ground between aesthetic and rational intuitionism, or an eclipsing of the debate altogether, one thing is certain: his main concern is to demonstrate that we possess a moral faculty which motivates our actions and, above all, claims authority over the other principles in human nature. A. E. Taylor is of like mind:

> What is Butler's real position on the questions which must be answered differently by a "rationalist" and by a "sentimentalist" in ethics? In the *Sermons* themselves he avoids taking up any definite position, and, as I have said, I think him justified on the ground that, for his

> immediate purpose, which is simply to insist on the authority of the moral law, a discussion of the precise way in which that law is apprehended would be irrelevant and disturbing. (1926, p. 278)

c) Conscience as Fallible

How reliable are the dictates of our conscience, according to Joseph Butler? Can conscience mislead someone desirous of acting virtuously?

Conscience operates when the fair, honest person, unsure as to her duty in a particular situation, asks herself what she should do, thus triggering "the rule of right within" (Sermon III, (3), p. 63). Conscience informs us of our duty and assists us in the discharge of it. In Sermon III, Butler states that we have the moral law within ourselves and that its dictates are almost always clear. In Sermon VII, *Upon the Character of Balaam*, he says: "In all common ordinary cases we see intuitively at first view what is our duty, what is the honest part." ((14), p. 117). For Butler, most doubt and deliberation with regard to duty is an endeavouring to explain it away, and thus deceive ourselves:

> The inquiries which have been made by men of leisure after some general rule, the conformity to, or disagreement from which, should denominate our actions good or evil, are in many respects of great service. Yet let any plain honest man, before he engages in any course of action, ask himself, Is this I am going about right, or is it wrong? Is it good, or is it evil? I do not in the least doubt but that this question would be answered agreeably to truth and virtue, by almost any fair man in almost any circumstance. (Sermon III, (4), p. 63)

The main exceptions Butler notes to the reliability of the dictates of our conscience concurring with the standards of virtue are self-deceit, superstition (indulgences or atonements) and self-partiality (dishonesty and unfairness of mind). Following conscience is thus no guarantee of virtue even though our

intention is invariably to act virtuously when we do so. This is due to our failing or incomplete moral character, which leads us more often than not to overlook rules we apply to others and which apply as well to the course of action in which we are engaged. In this regard, Duncan-Jones comments: "The authority of conscience, in Butler's full sense, will belong only to an enlightened conscience, conscience in the Kantian sense. For only the enlightened conscience will be responding to authentic moral truths, which have intrinsic stringency." (1952, p. 94). Because of our imperfect nature and ignorance, conscience, like speculative reason, is not infallible.

Béla Szabados addresses the issue of a corrupt conscience (the phenomenon of Eichmann) and questions whether the evolution of our moral character, on an individual and societal level, may influence the reliability of listening to our conscience in order to act virtuously. For Butler, virtue is a settled, habitual disposition or temper of mind. In the *Analogy*, he describes us as being in a state of probation, or trial, in the present life. We are led astray, as to our own interests as well as to those of others, by a wrong or immoral education, the ill behaviour of others (be it viciousness or simply a bad example), dishonest artifices, superstitions or mistaken notions of common opinion, and habits of indulgence. Thus, wrong behavior in youth often increases the difficulty of right behaviour in mature age (pp. 109-110). At the same time, Butler believes that we are capable of moral improvement through discipline, by acquiring habits of virtue, *i.e.* "by recollecting the practical impressions which example and experience have made upon us: and, instead of following humour and mere inclination, by continually attending to the equity and right of the case, in whatever we are engaged, be it in greater or less matters; and accustoming ourselves always to act upon it" (p. 123).

Szabados concludes, along with Butler, that a conscience continually ignored can lead to its corruption and unreliable authority: "by these means conscience may be laid asleep, and they may go on in a course of wickedness

with less disturbance." (Sermon VII, (10), pp. 114-115)[41]. However, ultimately, we are responsible for the results of such an aborted conscience, having chosen the path of vice over virtue. Szabados explains:

> It seems to me that Butler provides arguments which seem to show that his views are consistent. His view is this: (1) to have an errant conscience is something that a person is morally responsible for (unless it is due to "unconquerable" ignorance of fact); (2) to discern an errant conscience requires a conscientious act which itself requires the capacities involved in having a conscience. Thus conscience has ultimate moral authority. (1976, p. 465)

In other words, for Butler, it is the moral faculty itself which enables us to judge that we have erred from the path it dictates. Conscience's authority is revealed to us through its exercise. Ignored and put aside, unused, conscience may eventually become so rusty that it can no longer fulfill its purpose, but properly maintained and employed, or where possible revived, it will direct us to virtue.

In conclusion, Butler's moral conscience is a double-tiered faculty. On one level, conscience approves and disapproves of specific instances of virtue or vice as it reflects upon actions, motives and characters. Conscience can also be used to deduce certain universal standards of virtue or principles of conduct. On the other level, conscience approves of itself upon reflection as the faculty which has moral authority over all other principles in human nature. It is this second-order approbation which reveals to us the supreme authority of moral conscience.

---

[41] Gertrude Elizabeth Margaret Anscombe believes that "Butler exalts conscience, but appears ignorant that a man's conscience may tell him to do the vilest things." (1954, in 1981, p. 27)

Now that we have completed an overview of the nature of the moral faculty, let us examine how conscience derives it authority in Butler's doctrine of virtue.

3.2 Butler's Doctrine of Virtue

a) The Follow Nature Doctrine

In his Preface to the *Sermons*, Joseph Butler introduces the two premises of his doctrine of virtue: 1) that our nature is adapted to virtue; and 2) that virtue consists in following nature. Throughout Butler's discussion of virtue and human nature, these two premises intertwine to found his notion of moral agency, which ultimately rests on the authority of conscience.

According to Butler, testimony to the fact that *our nature is adapted to the virtuous life* is found in our capacity for moral agency due to the possession of conscience. But how then are we virtuous? When Butler says that *virtue consists in following nature*, he is replying that we are virtuous when we attribute to the moral faculty, the rank and function for which it was designed. The possession of conscience is thus necessary but not sufficient to make us virtuous: we must consider the use we make of it. For Butler, the first premise, that *our nature is adapted to virtue*, is logically and factually prior to the second, *that virtue consists in following nature.* Initially, we study the world around us in order to ascertain the ultimate purpose of our life in answer to the question: "How should we live?". We discover that we are adapted to virtue, since we have a supreme and authoritative moral faculty, conscience; following which we ask ourselves what we must do in order to live virtuously. Once again, we examine the external world and our inner subjectivity, and find that we can attain virtuous living only by according due regard or proportion to the

moral faculty as well as to the other principles in the system of human nature. Let us now deal with the two premises separately.

(i) Our nature is adapted to virtue

In Butler's view, nature, and more specifically human nature, is a system, with a purpose which can be deduced from the relation of its parts. Just as a man-made or artificial creation has such a purpose or utility, for example, a watch to tell time, all living beings have a purpose[42] which can be determined by a thorough examination of their nature. The primary purpose of living beings is to assure their survival, and this is carried out instinctually by animals as well as humans. Humans, however, are not ruled solely by instinct; our "constitution is put within our power", to use Butler's words. We may question Butler and ask why our purpose is not simply to act whimsically, spontaneously living as we choose. Butler's answer is, first of all, that human nature is organized as a system or constitution, which renders purposeful *versus* purely instinctual human activity possible. By studying the relation of the parts of the constitution of human nature to the whole, we can determine which type of activity we are naturally adapted to, and pursue it. Finally, the possession of a moral faculty, conscience, indicates that we are adapted to virtue. According to Butler, poverty, disease, disgrace, tortures and death are less counter to our nature than vice (*Sermons*, The Preface, (15), p. 11 and Sermon III, (2), pp. 61-62). While much in life may eat away at the human body and soul, depleting its vital energy, none are so damageable and thus "unnatural" as vice. Vice, by definition, is the undermining of the power over our constitution

---

[42] See Sermon III, (2), footnote 2, p. 62: "Thus the body is a *system* or *constitution*: so is a tree: so is every machine. Consider all the several parts of a tree without the natural respects they have to each other, and you have not at all the idea of a tree; but add these respects and this gives you the idea. The body may be impaired by sickness, a tree may decay, a machine be out of order, and yet the system and constitution of them not totally dissolved. There is plainly somewhat which answers to all this in the moral constitution of man."

which we were given, and the corruption of the highest authority in human nature, our conscience:

> It is from considering the relations which the several appetites and passions in the inward frame have to each other, and above all the supremacy of reflection or conscience, that we get the idea of the system or constitution of human nature. And from the idea itself it will as fully appear, that this our nature, *i.e.* constitution, is adapted to virtue, as from the idea of a watch it appears, that its nature, *i.e.* constitution or system, is adapted to measure time. What in fact or event commonly happens is nothing to this question. Every work of art is apt to be out of order: but this is so far from being according to its system, that let the disorder increase, and it will totally destroy it. (*Sermons*, The Preface, (14), pp. 10-11)

Thus, when Butler speaks of our "nature", he is not only referring to the different parts which constitute it, but includes the relationship between these parts in the constitution of human nature. As states Brian Hebblethwaite, Butler's notion of "nature" is normative, with virtue as the true end of our existence:

> For Butler does not offer his account of conscience and virtue in human nature in purely descriptive, naturalistic, terms. On the contrary, value is built into his account from the very start. We are not given a description of human nature in neutral, value-free, terms, from which evaluative conclusions are then supposedly drawn. The authority of conscience, on Butler's view, is an evaluative principle already built in to our nature. (in Cunliffe 1992, p. 205)

(ii) Virtue consists in following nature

In response to the question concerning what we must do in order to live virtuously, Butler gives the answer of the Stoics: that virtue consists in following our nature and vice in deviating from it (*Sermons*, The Preface, (13),

p. 7). However, "following nature" does not mean acting as any part of human nature may dictate, but rather in harmony with our entire constitution. Actions which are in conformity with human nature in this sense are virtuous, in contradistinction to actions which follow a single principle or propulsion, but are disproportionate in relation to our whole being. Shaftesbury's influence as well is evident here, as notes W. R. Matthews:

> The characteristic property of the good is to be *harmonious* or *proportionate*. In applying this general principle Shaftesbury makes use of the idea of a *system*, which Butler afterwards employed with such effect. Anything is good which is in harmonious relations with the system of which it forms a part. Hence it follows that human goodness consists in being in harmony with the species of which the individual is a member, and virtuous conduct is that which conduces to the good of the species as a whole. [...] The idea that the Good is harmonious is applied also to the individual. The virtuous man is one who maintains a "balance" between "self-affections" and "natural affections", which is Shaftesbury's term for the social and altruistic impulses. (1967, pp. xx-xxi)

Moreover, for Butler, following nature in the normative sense cannot mean simply acting as we please, for if this were the case, speaking of deviating from nature would be absurd, and following nature would have no meaning whatsoever: everything would be natural and it would be ridiculous to use nature as a guide to moral action. All principles are "natural" in the weak sense of the word. The passions and affections are as necessary a guide to conduct as the reflective principles of self-love and benevolence. However, following nature does not represent acting on the propulsion which happens to be the strongest at the time. This would be a violation of our true nature, says Butler, and would result in patricide being as "natural" as filial duty, which is a proposition that we definitely do not endorse (Sermon II, (17), pp. 58-59).

For our actions to be "natural" according to Butler, each principle must be attributed its proper place in the scheme of human nature, the reflective principles of self-love and benevolence regulating the passions. However, for our actions to be "in the highest and most proper sense, natural or unnatural"[43], the faculty which is in kind supreme over all other principles, and which bears authority due to the fact that it is this principle which makes man a law unto himself, should be our proper governor and direct all other principles:

> Whoever will consider his own nature, will see that the several appetites, passions and particular affections have different respects amongst themselves. They are restraints upon, and are in a proportion to each other. This proportion is just and perfect, when all those under principles are perfectly coincident with conscience, so far as their nature permits, and in all cases under its absolute and entire direction. [...] so far as this superiority is maintained, the character, the man, is good, worthy, virtuous. (Sermon III, (2), footnote 2, pp. 62-63)

In Butler's view, when we accord due proportion to the principles in our nature by letting our conscience guide self-love and benevolence, and the two latter reflective principles reign over the passions, we bring ourselves closer to the virtuous life. In so respecting the unity of our human nature, we are less divided and we gain personal fulfilment, as well as the peace of mind or happiness which accompanies living the law of our nature. This, for Butler, is the true meaning of "Reverence thyself" (*Sermons*, The Preface, (25), pp. 14-15).

At this point, we would like to examine two critiques of Butler's "follow nature" doctrine. The first is the "Full Naturalistic Thesis" which

---

[43] "This *prerogative*, this *natural supremacy*, of the faculty which surveys, approves or disapproves the several affections of our mind and actions of our lives, being that by which men *are a law to themselves*, their conformity or disobedience to which law of our nature renders their actions, in the highest and most proper sense, natural or unnatural" (Sermon II, (9), p. 54).

Nicholas Sturgeon attributes to Butler and which, according to the former, renders Butler's doctrine of the supremacy of conscience superfluous. The second accuses Butler of a tautology or vicious circle in his definition of virtue, and is formulated by Henry Sidgwick and Leslie Stephen.

b) Critiques of the Follow Nature Doctrine

Nicholas L. Sturgeon's "Full Naturalistic Thesis", which he believes Butler is committed to, is "that conscience never *favors* or *opposes* any action, except on grounds which include its naturalness or unnaturalness", or if one prefers, its conformity to the nature of the agent (Sturgeon 1976, p. 328). The basis for Sturgeon's thesis are two propositions which he attributes to Butler: "(1) that virtue consists in following nature, vice in deviating from it, and (2) that whenever conscience approves or disapproves of an action, it does so on grounds of the virtue or vice of the action." (p. 325).

Sturgeon further argues that, according to Butler, it is the opposition of a superior principle that makes an action unnatural, and the favour of a superior principle that makes an action natural. It is important to note here that the superior principle need not be conscience, but could be self-love, or even benevolence if neither conscience nor self-love opposes the action: "On the assumption that conscience and self-love do not conflict, and that these are the two highest principles in human nature, the opposition of either will indeed be sufficient to guarantee that an action is unnatural, even if conscience is the superior of the two." (pp. 339-340). Sturgeon ranks benevolence below conscience and self-love: "For only in those cases in which an action is neither opposed by self-love nor exclusively favored by it does it accord to benevolence any bearing whatever on the naturalness or unnaturalness of the action." (p. 342).

Sturgeon finally concludes that conscience favors or opposes an action, ultimately, only because another superior principle (*i.e.* other than conscience)

favors or opposes it first. In his words, conscience "cannot operate in a vacuum" any more than self-love can without the passions (p. 330, quoting Tom Aerwyn Roberts 1973 on self-love, at p. 51). For the reason that conscience thus "charts no independent course of its own", its authority is superfluous (p. 345): "For purposes of determining the naturalness or unnaturalness of actions, then, not only the doctrine that conscience is uniquely supreme, but even the doctrine that it is a superior principle at all, is entirely superfluous." (p. 347).

In response to Sturgeon's claim that Butler's doctrine of the supremacy and superiority of conscience is superfluous, it should be noted that Butler states that recognition of the supreme authority of conscience is at the core of our capacity for moral agency and of our understanding of what it means to live virtuously, or conformably to the constitution of human nature. This is the essence of his theory that our nature is adapted to virtue and that we should therefore "follow nature". Let us examine this thesis and Sturgeon's arguments more closely.

First of all, Butler does not seem committed to the "Full Naturalistic Thesis" that Sturgeon attributes to him (*i.e.* that "conscience never *favors* or *opposes* any action, except on grounds which include its naturalness or unnaturalness"). Butler does not base the approval or disapproval of conscience on the naturalness or unnaturalness of an action, *i.e.* its conformity to the nature of the agent as a whole, in the constitutional sense of the term. The grounds of reflex approbation or disapprobation are the virtuous or vicious tendency of the action (after consideration of its intention and not the actual consequences), *i.e.* its respecting or contravening universal standards of morality such as justice, veracity, prudence and regard to common good:

> Nor is it at all doubtful in the general, what course of action this faculty, or practical discerning power within us, approves and what it disapproves. For, as much as

> it has been disputed wherein virtue consists, or whatever ground for doubt there may be about particulars; yet, in general, there is in reality an universally acknowledged standard of it. It is that, which all ages and all countries have made profession of in public: it is that, which every man you meet puts on the show of: it is that, which the primary and fundamental laws of all civil constitutions over the face of the earth make it their business and endeavour to enforce the practice of upon mankind: namely, justice, veracity and regard to common good. (*Dissertation*, (1), p. 248; see the virtue of prudence as well at (6)-(7), pp. 251-253)

With all deference to Sturgeon, Butler's conscience does not operate in a vacuum; it has the superior principles of reflection, *i.e.* self-love and benevolence, as well as the particular passions to employ. However, its grounds for favoring or opposing an action motivated by one of these principles go above and beyond them, and concern the realm of morality. A "virtuous" action is "natural" in the sense that it follows the supreme principle in the hierarchy of human nature, moral conscience, and thus respects the constitution of human nature; just as a "vicious" action is "unnatural" in that it follows the lower, albeit stronger principles, in opposition to the dictates of conscience. However, the two epithets are not synonymous for Butler: in order for an action to be qualified as "virtuous" and not just "natural", as when a superior principle masters the passions, its intent must respect certain moral standards. Conscience is a guide to virtuous conduct, directing us to standards of right behaviour. As Terence Penelhum notes:

> I might well act naturally by doing what conscience tells me to do, and act unnaturally by doing what conscience tells me not to do; but this does not show that conscience tells me to do things because they are natural, or not to do them because they are unnatural. On the face of it, it tells me to do them because they are *right*, and not to do them because they are *wrong*. In

> view of a recent criticism of Butler, this is of great importance. (1985, p. 22)

How, then, are we to understand Butler's "follow nature" doctrine? Sturgeon's two propositions, *i.e.* 1) that virtue consists in following nature, and 2) that conscience approves or disapproves of an action on grounds of the virtue or vice of the action, can indisputably be attributed to Butler. Sturgeon combines the two propositions to end up with "conscience never *favors* or *opposes* any action, except on grounds which include its naturalness or unnaturalness". However, both propositions are in fact concerned with Butler's demonstration of conscience as the moral faculty, approving virtue and disapproving vice. Butler is not attempting a definition of virtue nor fixing the content of virtue when he refers to nature. Rather, he is describing the proper functioning of human nature, or how we should act in order to fulfil our ultimate purpose in life, virtuous living. Millar thus differentiates between a *guide* to virtue and a *standard* of virtue: "A guide to virtue is not necessarily a standard of virtue. Conformity with the standard of virtue is what makes an action, quality or mode of life virtuous. To say that our nature points to that mode of life which is virtuous is not yet to say that our nature is the standard of virtue." (1988, p. 172).

The answer Butler's "follow nature" doctrine gives to those wishing to act virtuously but wondering how, or questioning why they should act virtuously in the first place, is the following: act conformably to the law of your nature, which you have a natural obligation to follow, by pursuing that which is within your power, *i.e.* by ensuring that your moral conscience guides all principles, and that the superior reflective principles of benevolence and self-love direct the passions. This is Penelhum's explanation:

> I have said that Butler's claim that virtue consists in following nature is not intended to help identify what virtue requires, but to help ensure its practice. It is also not a conceptual thesis about what the word "virtue"

> means, though in developing it Butler makes comments about our ethical vocabulary. Butler claims that those who practise virtue are acting in the way that is consonant with the makeup given them by providence, and that this is one which includes the recognition of the supremacy of conscience. (1985, p. 61)

Darwall agrees with Penelhum that Butler's claim that virtue consists in following nature "is addressed to the 'obligation' to virtue, to whether there is reason to be virtuous." (in Cunliffe 1992, footnote 23, p. 216).

While Sturgeon is correct in saying that the opposition or favour of a superior principle "makes" an action unnatural or natural, in that in Butler's constitutional sense of the word it is "unnatural" to favour a lower principle over a higher one and "natural" to favour a superior principle to one beneath it, he misses the mark when he concludes that conscience favors or opposes an action because another superior principle favors or opposes it.[44] The "unnaturalness" of an act in the teleological sense is derived from a comparison between the action's intent and the "nature" or capacities of the agent, and as Butler makes clear, we are moral agents:

> Our perception of vice and ill desert arises from, and is the result of, a comparison of actions with the nature and capacities of the agent. [...] And this determination must arise from such comparison, and be the result of it; because such neglect would not be vicious in creatures of other natures and capacities, as brutes. [...] And hence arises a proper application of the epithets, *incongruous*, *unsuitable*, *disproportionate*, *unfit*, to

---

[44] See Darwall in Cunliffe 1992 at p. 220: "Above I said that Butler distinguishes between conscience's approval of an action and there being an obligation to follow this 'rule of right within'; and that his argument that in violating conscience we go counter to a law of our nature and thereby act unnaturally is addressed to the latter. The Full Naturalistic Thesis ignores this distinction. It maintains that the grounds that constitute conscience's authority, namely considerations of natural superiority and inferiority, are the very grounds on which conscience initially approves or disapproves conduct."

> actions which our moral faculty determines to be vicious. (*Dissertation*, (5), pp. 250-251)[45]

> Therefore instead of the words *disproportionate to his nature*, the word *unnatural* may now be put; this being more familiar to us: but let it be observed, that it stands for the same thing precisely. (Sermon II, (10), p. 55)

Conscience is not a formal, symbolic, vacuous feature in Butler's system of human nature, it is the moral faculty par excellence which determines whether actions are virtuous or vicious. Sturgeon never fully discusses nor settles the question of which superior principle (besides conscience) judges a motive to be virtuous or vicious, although he deems this to be a requirement in determining the virtue or vice of an action. We submit that, according to Butler, only the inner voice of conscience could make such a determination. How could conscience then be superfluous in moral decision-making? Conscience, the supreme principle in Butler's theory, is only concerned with the "naturalness" of actions to the extent that we equate "naturalness" with "in accordance with virtue". No other superior principle of reflection, be it self-love or benevolence, is capable of discerning the "naturalness" of actions in this last sense, that of moral fitness. In Millar's terms:

> Butler clearly assumes in this passage that, by and large, we all know what is morally right. This is implied in his thinking that we have a faculty, conscience, which is a 'practical discerning power', imbued with standards of 'justice, veracity, and regard to the common good' which are 'universally acknowledged'. There is no suggestion here that conscience works by looking to our nature as the standard of right. The standard is rather written into our hearts and minds in that we judge and feel in accordance with it. (1988, pp. 171-2)

---

[45] See Millar 1988 at p. 178: "it is perfectly possible on Butler's principles to explain the sense in which virtue consists in following nature, and vice in deviating from it, *without* recourse to the natural hierarchy of motivating principles. It is the notion of adaptation which bears the explanatory load, not the hierarchy of principles."

Nature is a guide to virtuous behaviour, to be followed through the respect of the superiority of its principles and above all the supremacy of conscience, but it is not the standard of virtue. Similarly, moral conscience indicates the way to virtue, but the substance of virtue is found in the timeless standards of morality which are engraved in our hearts and which are ours to discover, since we are a law unto ourselves. Conscience, the "guide of life", could be illustrated by the metaphor of the lighthouse: from the shore, it purveys the seas, and brings to our attention a variety of phenomena. However, the purpose for which it was created is to serve as a beacon to ships, in all types of weather, both to bring them to shore, and to avoid them from crashing against it. Conscience is the lighthouse of our soul which accompanies us through life.

According to certain commentators, such as Sir Leslie Stephen, Butler's reasoning behind his doctrine of the supremacy of conscience and that virtue consists in following nature involves a vicious circle, making conscience ultimately a "self-evidencing power":

> We disapprove immoral actions, and immoral actions are those which we disapprove. What then is this special supremacy of conscience? Why is it exceptional? [...] The conception of a self-evidencing power seems to involve a vicious circle. (1962, vol. II, 51., p. 42)[46]

In Stephen's view, conscience must, in some way, derive its credentials from another authority than itself, and he believes that for Butler, this authority is God's will (p. 43). Lefevre, however, offers an alternative explanation for Stephen's dilemma: "the proposition that our nature is adapted to virtue [...] is

---

[46] See Sturgeon 1976, footnote 14, p. 332; and Sidgwick 1981, at p. 81: "Nor does it help us to say that the supremacy of Reason is Natural, as we have started by assuming that what Reason prescribes is conformity to Nature, and thus our line of thought would become circular: the Nature that we are to follow must be distinguished from our Practical Reason, if it is to become a guide to it."

a self-evident deduction from our structure" because "it is only when we take into account the supremacy and authority of conscience that we get the idea of the constitution of human nature" at all (1899, p. 131). Butler's conscience is indeed, as Darwall puts it, "self-authorizing" (1995, p. 244 *ss.*).

According to Sidgwick, Butler's conscience is an "arbitrary authority", which informs us of our obligation to virtue, and at the same time "empties virtue of all practical content"[47] by making it the sole good, thus referring us back to conscience itself:

> Butler assumes with his opponents that it is reasonable to live according to Nature, and argues that Conscience or the faculty that imposes moral rules is naturally supreme in man. It is therefore reasonable to obey Conscience. But are the rules that Conscience lays down merely known to us as the dictates of arbitrary authority, and not as in themselves reasonable? [...] But if Conscience is, after all, Reason applied to Practice, then Butler's argument seems to bend itself into the old circle: 'it is reasonable to live according to Nature, and it is natural to live according to Reason'. (1981, p. 378)

In response to Sidgwick, though Butler's reasoning may appear circular, this is due to the way conscience claims authority in the moral sphere: in summoning us to follow its dictates and our nature, it not only reveals to us our duty, but our moral nature as well and, ultimately, the supreme authority that the moral law has over this nature.

This raises two interesting questions: Must conscience have a foundation for its authority, and why? If so, what is this foundation in Butler's moral theory? In response to these questions, we will now examine the arguments for the authority of conscience.

---

[47] Sidgwick 1981, at p. 377 comments on the Stoic formula of "Life according to Nature": "But if Virtue is thus declared to be a science that has no object except itself, the notion is inevitably emptied of all practical content."

CHAPTER 4

Arguments for the Authority of Moral Conscience

In this chapter, we will examine three arguments for the authority of conscience attributed to Butler. The first argument is "teleological/functional" and concerns the purpose to which our nature is adapted, *i.e.* virtue. The second is "constitutional" and reflects the relative ranking between the different principles in the hierarchical system of human nature, with conscience as supreme. Finally, the third is called "autonomist" and relates to the moral autonomy of the will which is evidenced by the exercise of conscience. In conclusion and throughout the chapter, we will consider whether Butler's theory of moral conscience stands its ground without its theological backing.

4.1 The Teleological/Functional Argument

Butler's initial argument for the authority of moral conscience is illustrated by the analogy of a watch, created to measure time. Butler submits that if we study our nature, we will observe that it functions properly only when we fulfill the apparent end for which we were designed, *i.e.* virtue. Vice is a malfunction, where our system is out of order. Since our purpose is to pursue virtue, and access to the standard of good and right is obtained through our moral conscience, this principle is authoritative in human nature. Thus, it would be absurd to put parricide and filial duty on an equal footing.

The theological foundation for this teleological argument is that the Author of Nature designed us for the virtuous life (God being good and having created us in his image). However, since we possess free will, we may choose to disrespect this path and pursue vice; inevitably, this will lead us to ruin, just as if we had attempted to use a watch as a nutcracker. A further guarantee of the unhappiness of vice is God's being the moral as well as natural governor of

the universe, and the knowledge that we will receive our just reward and punishment in the afterlife if not in this life.

According to Lefevre, the validity of Butler's teleological argument rests on his assumption that there are no vicious principles *per se* in human nature:

> If it had been found true that within human nature there were contradictory principles leading to opposing ends, then either it would have been illogical to deduce virtue as the end from the structure of our being, or our nature could not be regarded as constituted by organically connected parts. (1899, p. 140)

How convincing is Butler's claim that we are "adapted to virtue"? Why not, for example, happiness? Interestingly, Butler's posits that though God himself may be a utilitarian, the production of happiness is not our ultimate end.[48] Due to our ignorance, we can never know what will truly be to our happiness, whereas the voice of conscience speaks to us with the certainty of moral obligation. Since we can never be certain that vice will be to our interest and the moral law presents itself to us with authority, we can only gain in following the voice of our conscience. In the long run, Butler believes that duty and interest coincide and that it is impossible vice can be to our advantage in this world: "It will immediately appear, that vice cannot be the happiness, but must upon the whole be the misery, of such a creature as man; a moral, an accountable agent." (*Upon the Character of Balaam*, (16), p. 119).

In Millar's opinion, Butler's teleological argument is far from convincing:

> Butler looks at the facts through the lens of his theology, that is to say, with the aid of background

[48] *Dissertation*, (8), p. 255: "And therefore, were the Author of Nature to propose nothing to himself as an end but the production of happiness, were his moral character merely that of benevolence; yet ours is not so."

> beliefs about how human beings would be designed if created by God. These background beliefs provide a positive reason to expect that we are made for virtue and not for vice and thus a reason to seek an explanation of appearances to the contrary. Thus, for Butler, the fact that most human beings fall far short of the balanced life which would be led by the truly virtuous does not count against his basic claim that we are, nevertheless, suited to virtue. I do not wish to argue here that Butler is wrong, only to point out that if we eschew his theology it will be a serious question whether Butler's optimistic conception of our nature can be sustained. (1988, p. 183)

Although Butler's teleology has the theological underpinnings we have just described, is his so-called "optimism"[49] about human nature by that means necessarily religious? It would seem to concur with common sense that we do not exist in order to do evil in this world. When we wreak havoc and destruction, it is usually in the name of some greater good. Of course, there are a multitude of ends which we may set for ourselves, but ultimately, most philosophers and laymen agree that they come down to a select few, such as happiness or pleasure, the good of self and others, or self-realization. An evolutionary theory may prefer survival of the species or self-preservation. The question is whether it is through a moral faculty such as Butler's conscience that we have access to such ultimate end(s). Could it not be given simply through reason or, rather, the senses?

We concur with Butler that the inklings we have to the nature of our purpose in life are through the moral faculty. Reason may help us structure our intuitions and the senses feed them with impressions, but, in the end, our direction is relegated to moral decision-making. "What *should* I do with my

---

[49] See Basil Willey 1949 at pp. 77-78 for a contrary view: "In passing from Shaftesbury to Butler we pass from an optimistic to a relatively pessimistic theory of the world. It is not strange that the champion of orthodoxy should be more pessimistic than the heretic, for the Christian tradition had always been associated rather with a sense of the imperfection of

life?", whether there is a God or not, is a moral question and the "should" carries with it the belief that there is indeed such a purpose for which we were created.

Millar concludes that Butler's "core concept of adaptation" of human nature relies on his theistic teleology of "intelligent design and craftsmanship": "The watch's constitution is explained by the fact that a watchmaker has made it so that it will be suitable for measuring time. The constitution of our nature is explained by the fact that we are made by God for virtue." (1992, p. 488). Millar then offers an alternative teleological explanation, that of natural selection in evolutionary theory. Sober and Wilson as well use "adaptationist methodology" in their evolutionary, psychological and philosophical discussion of egoism, hedonism and altruism. They describe "adaptationism" as "a method for investigating nature. This is the idea that a useful procedure for studying an organism is to ask, 'What would the organism be like if it were well adapted to its environment?' Posing this question does not commit one to the position that the organism actually *is* well adapted." (1998, p. 11). We submit that Butler uses a similar methodological approach in investigating human nature in his moral theory. The conclusion he arrives at is that we are adapted to virtue, since we approve of universal moral standards such as justice, veracity, prudence and regard to common good. Alan Brinton seems to be of like mind:

> I conclude that Butler does not really have the problem Millar thinks he has. Butler finds within human nature indications that we are adapted to a life of justice, veracity, and doing good to others. There is no reason to think that he is not perfectly satisfied with probability in this case, as he is on other important matters. He does not, therefore, have to invoke some kind of guarantee from God; his ethical theory does not

---

Nature and of man, in their present state, than with any optimism of the eighteenth century type."

> require divine providence in the way Millar thinks it does. (1991, pp. 329-330)

In other words, Butler believes that a careful study of human nature will reveal that we are well adapted to virtuous living; and we can either help or hinder the realization of our life's purpose by choosing to follow the dictates of our conscience or not.

Let us now examine the constitutional argument for the authority of conscience.

4.2 The Constitutional Argument

In Butler's system of human nature, the parts or principles are related to one another in a manner which is conducive to the realization of the end for which we are adapted, *i.e.* the virtuous life. Following our nature in the normative sense, then, means allowing our supreme internal principle, the one that claims authority, to govern. By definition, this principle is the one that indicates to us the path of virtue, and not some other end, however appealing that end may be. We have seen that, for Butler, certain principles in the hierarchy of human nature are superior to others in kind, although maybe not in strength or degree. He illustrates this with the difference between power and authority to rule under a constitution.

The particularity of conscience in what Darwall calls Butler's "self-regulated constitutional order" is that it claims title to rule (1995, pp. 275-6). When we follow an inferior principle rather than our conscience, we violate our constitution:

> The ruling metaphor is that of a *constitutional order*, which exists only if there are relations of *authority* – a truth about who is to govern when the claims of political actors conflict. Likewise, our internal nature

> forms a constitutional order only if there is a truth about which principles should govern when they conflict. In these terms, in arguing that disobeying conscience is unnatural, Butler is maintaining that it is *unconstitutional* – it "*violates*", as he sometimes says, our internal constitution. [...] Claim to a constitutional role is part of the very concept of conscience. It alone, from its very nature, claims authority and superintendency. (1995, pp. 261-262)

This disproportion to our nature as a whole occurs not because of the actual consequences of our action, nor due to the nature of the act considered in itself, but by "*comparison* of it with the nature of the agent" (Sermon II, (10), pp. 54-55). Butler reminds us that we are agents, and that our constitution is within our power. We are hence responsible for our acts and morally accountable. When we ignore this distinction by disregarding conscience, we discredit the principle which brings us totality or systemic and organic unity, and become like a watch out of order or a broken-down machine. As Lefevre summarizes:

> Assuming that from the inward frame of man and its natural adaptations we can ascertain what course of life and behavior that real nature points out and leads to, Butler argues from the fact of the existence and nature of conscience to the proper end of our being. He finds that, as the moral faculty, it is designed for and hence adapted to virtue. Since it is not only the supreme part, but also the synthetic principle of the human organism, its goal becomes identical with the complete end of man, or, in other words, obedience to conscience secures the realization of man's whole nature. (1899, p. 134)

Conscience's authority is thus doubly guaranteed: by its role in fulfilling the true end to which we are adapted, virtue, and revealing our duty to us (the teleological argument), and by its capacity to obligate in claiming governance over the other principles in human nature (the constitutional

argument)[50]. Let us now follow with the third and final argument for the authority of conscience, namely, that of moral autonomy of the will.

4.3 The Autonomist Argument

What we call the "autonomist" argument for the authority of conscience is Butler's suggestion that conscience is the faculty which makes moral autonomy possible, *i.e.* "that by which men *are a law to themselves*" (Sermon II, (9), p. 54), since it is through the possession and exercise of conscience that we arrive at binding moral judgments. This is the third sense of its authority. Conscience provides us with moral reasons and, by the same token, let us know that such reasons are preponderant. In so doing, it notifies us of our obligation to act as it dictates, and further motivates us to do so through this very realization of our moral agency:

> But allowing that mankind hath the rule of right within himself, yet it may be asked, "What obligations are we under to attend to and follow it?" I answer: it has been proved that man by his nature is a law to himself, without the particular distinct consideration of the positive sanctions of that law; the rewards and punishments which we feel, and those which from the light of reason we have ground to believe, are annexed to it. The question then carries its own answer along with it. Your obligation to obey this law, is its being the law of your nature. That your conscience approves of and attests to such a course of action, is itself alone an obligation. Conscience does not only offer itself to show us the way we should walk in, but it likewise

[50] Brinton names the first "material virtue" and the second "constitutional virtue": "There are, I conclude, two quite different senses of 'virtue' in Butler's moral philosophy: the 'consists in' sense, which we may call 'constitutional virtue', and the 'adapted to' sense, which we may call 'material virtue'. To say that an action is constitutionally virtuous is to say that it is in accord with the rightfully governing faculty of the mind, namely conscience. To say that an action is *materially* virtuous, on the other hand, is to say that it is in accord with principles such as 'justice, veracity and the common good'." (1991, p. 328)

> carries its own authority with it, that it is our natural guide; [...] (our emphasis) (Sermon III, (5), p. 64)

In other words, our conscience gives us *conclusive* (Broad 1930, p. 78) or *overriding* reasons[51] to do what is right. In essence, conscience reveals to us through its authority that there are no better reasons than moral ones. This is what A. E. Taylor calls the "intrinsic authority" of the moral law (1926, p. 299) and Duncan-Jones the "intrinsic stringency" of moral qualities which give "prepotent" reasons for acting: "Acknowledging a moral quality and doubting whether there is any reason for acting in a certain way towards what possesses it is rather like giving the proof and doubting the conclusion." (1952, p. 86). Only another moral reason can outweigh a moral reason. Note that Darwall finds the term "conclusive reason" to be misleading when applied to Butler's conscience, and stipulates that the term must be given a *de jure* and not a *de facto* reading[52]. According to Darwall, in Butler's system of morality, the practical reasoning realizes autonomy in the following way:

> Normative practical notions, [Butler] argues, are intrinsic to autonomous *agency*; Butler holds a *normative theory of the will*. Agents can *regulate* or *govern* themselves only if they act on a normative conception they *accept*. But neither, like Clarke, does Butler think that moral *obligation* derives from a normative fact whose existence is independent of the form of practical thinking that makes autonomous agency possible. Rather moral obligation consists in its being the case that an agent would *regard* herself as having conclusive motive de jure – that is, conclusive *reason*, were she to exercise the form of practical thought that makes autonomous action possible. (1995, p. 248)

---

[51] See Robert L. Arrington 1998 at p. 222: "Its judgments are experienced by us as overriding, as outweighing all other considerations." and Darwall 1995 at p. 325.

[52] Lefevre shares this view: "And as the moral nature is the highest and distinctive part of man, that which makes human nature a constitutional whole, the faculty which, as a matter of fact, does pass moral judgments, is, both *de facto* and *de jure*, supreme, and lays upon us the most intimate obligation." (1900, p. 405).

However, although, for Butler, moral *obligation* derives from the discovery by the agent of her autonomy, and not from an independent normative fact as per Samuel Clarke, conscience does not create the moral law, it only constitutes moral agency. Lefevre remarks: "Conscience is simply the *capacity* for virtue. It does not *make* morality, but it makes moral action possible." "But conscience no more creates morality than the eye creates the things it sees, or the feeling of shame that which is its ground and cause and explanation, or, to use another of Butler's illustrations, than the watch creates the time it measures." (1900, p. 401)[53]. Conscience is simply the faculty through which we have access to the moral order of universal standards, virtuous principles and fitness of actions which guide us.

There are different criticisms of this third argument for the authority of conscience. Firstly, there is the consideration that even if conscience, due to its authority, delivers overriding reasons for action, this cannot serve as justification for the supremacy of conscience, or else conscience becomes self-justifying. We are back to the problem of circular argumentation. W.A. Spooner's answer to the apparent circularity of Butler's argument is the following: it is "the very nature of all ultimate and immediate judgments" "that we can give no further reason for the verdict" they pronounce (1901, pp. 112-113)[54].

Wendell O'Brien, while mentioning this difficulty, proposes another explanation: "It is of course possible that Butler has some notion of 'matter of fact' which is such that what we ought to do and what the best reasons for

---

[53] "Since then our inward feelings, and the perceptions we receive from our external senses are equally real; to argue from the former to life and conduct is as little liable to exception, as to argue from the latter to absolute speculative truth. A man can as little doubt whether his eyes were given him to see with, as he can doubt of the truth of the science of *optics* deduced from ocular experiments. And allowing the inward feeling, shame; a man can as little doubt whether it was given him to prevent his doing shameful actions, as he can doubt whether his eyes were given him to guide his steps." (Sermon II, (1), pp. 48-49).

[54] "–is it not the very condition of such judgments to be exposed to the reproach of being arguments in a circle?" (Spooner, p. 112)

action are are questions of fact, and that we are attributing some fact/value distinction to him which he does not in fact recognize." (1991, p. 46). O'Brien then goes on to deny, in our view quite correctly, that Butler has a similar view of obligation.

However, Duncan-Jones, commenting on the nature of moral qualities attributed to actions, notes that many modern philosophers would indeed interpret the intrinsic stringency of conscience's edicts as simply an emotive statement or one expressive of our attitudes and habits (1952, p. 89-90). Jeffner, as well, believes that Butler appears to assume that the psychological authority of conscience (it approving of itself upon reflection) is in some sense a reason for its moral authority (1966, p. 213). Along the same line, W. D. Falk offers a psychological explanation of conscience whereby we internalize its authority: "When called upon to decide here and now which act to do they consider which act they have from within themselves the most conclusive reasons for doing; and they then no longer think of being morally bound as an external requirement, but as an ideally inescapable inner entanglement, a dictate of conscience." (1948, p. 137) These, however, are rather critiques of Butler's choice of naturalistic methodology in his demonstration of the authority of conscience, than proofs against its authority. As Lefevre argues:

> If, however, one remembers that Butler's psychological investigation is merely a method by which he sets out to ascertain the facts of human nature in order that an idea of the goal of that nature may be thence inferred, the force of such a criticism is lost. [...] The investigation of the facts has, Butler thinks, of itself shown both the adaptation and the obligation to the pursuit of virtue as the complete end of man, an end which appeals not merely to any one part of his nature, but to his nature as a consistent whole. (1900, p. 399)

The "problem" which Leites finds in Butler's ethics is that Butler asserts the obligation to obey conscience due to its natural authority, without

commenting on the certainty of our other moral obligations; he thus neglects to compare the certainties of the obligation to obey conscience and the obligation to promote personal advantage, as Sidgwick later attempted to do (Leites 1975, p. 50). Duncan-Jones formulates this as the comparison between "*there are* reasons" and "*I have* reasons" (1952, p. 88); and as we saw above, according to him, full authority of conscience belongs only to an "enlightened conscience" which will respond to authentic moral truths which have intrinsic stringency.

Darwall's answer and explanation of Butler's autonomist line is the following:

> It attempts a "transcendental deduction" of the authority of conscience as a necessary condition for the very possibility of a kind of *internal* order: autonomous agency or self-regulated internal constitution. [...] We could not even seriously raise the question of whether there is reason to follow conscience unless we had the capacity to answer it affirmatively and be guided practically by that answer. (1995, p. 329)

To conclude, even those convinced by Butler that the source of moral obligation and autonomy is the voice of our conscience, question the foundation of the moral standards, principles or duties which conscience reveals to us. Do they derive from evolving custom, social contract or biological adaptiveness or are they eternally immutable moral truths? Many believe that, for Butler, they originate in God's will.

Let us now consider whether the above arguments for the authority of conscience can be maintained independently of Butler's theology.

4.4 Independence of Butler's Ethics from his Theology

According to certain commentators, the foundation of Butler's theory of the authority of conscience is his theistic teleology, *i.e.* the Author of Nature's

design for us[55]. Without God's will and design, there is no guarantee that we were created for the purpose of virtue: "Butler's escape from the vicious circle really consists in his assumption that the conscience represents the will of God." (Stephen 1962, vol. II, 51., p. 43).

On the other hand, some writers claim that Butler's naturalistic ethics is not dependent on the belief in the existence of God or of an afterlife. They submit that although Butler was arguing primarily against the Deists, thus his insistence on God and Providence to conclude his arguments exclusive of revelation, Butler believed in "the moral fitness and unfitness of actions, prior to all will whatever" (*Analogy*, p. 284), similar to his teacher Samuel Clarke:

> But if it be intelligible to say, that *it is fit and reasonable for every one to consult his own happiness*, then *fitness of actions, or the right and reason of the case*, is an intelligible manner of speaking. [...] It doth not therefore appear, that moral right is any more relative to perception, than abstract truth is; or that it is any more improper, to speak of the fitness and rightness of actions and ends, as founded in the nature of things, than to speak of abstract truth, as thus founded. (*Analogy*, footnote at p. 143)[56]

Lefevre explains that Butler's teleological argument assumes that there are moral standards to be discovered in the same way as truth is apprehended:

> One could point out the same circle in regard to speculative reason that is said to exist in Butler's theory

---

[55] See, for example, O'Brien's conclusion in 1991, p. 55; Penelhum 1985 at p. 75; Michael S. Pritchard 1978 at p. 41, and Hebblethwaite: "Butler's ethic, I shall argue, has an essential religious foundation: It is not a divine command theory, but it *is* a form of religious natural law theory, with conscience possessing its supreme authority solely because that is how the Author of nature intended things to be." (in Cunliffe 1992, at p. 198).

[56] See Lefevre 1900 at p. 409: "Consequently, morality is in its last definition simply the 'eternal fitness' of things, and, like truth, is grounded in the nature of things."; and Silvan Solomon Tomkins 1934 at p. 47: "But if it is misleading to consider Butler's ethics in abstraction from his theology, it is no less violence to his thought to regard God as a *deus ex machina*, as Leslie Stephen does, for God and man alike are dependent upon the eternal fitness of things prior to will."

> of conscience. Truth is that which reason discerns to be true, and reason is the faculty which determines truth. As the escape from such a 'circle' lies in our discovery of the standards and tests by which we judge truth, so the escape from the same circle involved in the doctrine of conscience lies in our discovery of the standards and tests by which we judge morality. (1900, p. 400)

As Norton illustrates, the rational aspect of conscience and the exercise of reason in the speculative sphere are each in their own capacity ultimate in the sense that there is no higher authority in their respective spheres (1940, p. 95). However, whereas speculative reason is unreliable in its sphere, conscience is not.

According to Darwall, Butler's acceptance of Clarke's theory of eternal fitnesses, his teleological argument for the authority of conscience and his theology are in tension with his autonomist argument (1995, p. 248). Darwall also notes that Butler sometimes confuses the metaphors of the constitutional argument, that conscience has title to rule, with his theistic teleology, that conscience was designed by God to override. Darwall believes that the serious weakness in Butler's doctrine of the authority of conscience is that its commands are simply given:

> And what can guarantee convergence in conscientious judgments for every agent, especially if, as Butler sometimes suggests, human conscience's dictates depend on God's solution to a contingent problem of social engineering, namely, the problem of which conscientious dictates are likeliest to achieve the happiest whole? (1995, p. 330)

The content of Butler's conscience is thus "fixed contingently rather than by anything internal to the functioning of autonomous practical reason itself" (p. 330).

According to Duncan-Jones, Butler's ethical teaching is essentially non-theological, in that "it is possible to extract from Butler's writings a moral philosophy conceived in purely natural terms", where *good* and *right* are not analyzed in terms of God's will (1952, p. 142). However, as previously discussed, Butler would then "have had to admit that there is such a thing as "real ill-will", even if it is providentially overruled for a good end." (1952, p. 150). Furthermore, Duncan-Jones thinks that Butler is wrong that there is a universal standard of virtue across different ages and places, and that "the lack of agreement among moral philosophers themselves goes to refute the view that there is a distinct kind of moral law knowable by a distinct faculty." (1952, p. 166). Albert Edward Baker is of the same opinion:

> The conscience whose 'supremacy' and 'authority' Butler discusses is the conscience of the 'urbane', gentlemanlike, individual of the middle classes in the Church of England in 1730. The age was badly instructed with regard to other countries, different civilizations, and past ages. [...] Human nature is no longer the constant datum it was for eighteenth century thinkers. (Baker 1923, p. 113-114)[57]

These authors maintain that it is only since the theory of evolution that the moral sense is regarded as a product of development. We do not believe this to be historically correct. Butler consecrates an entire chapter in the *Analogy* to the question of moral discipline and improvement. He considered virtuous tendencies to be acquired by trial and error, practice and habit, with moral education forming the character. When Butler says of the universal standard of morality in the *Dissertation* that "It is that, which all ages and all countries have made profession of in public" ((1), p. 248), he is clearly addressing the issue of the reliability of the edicts of moral conscience across

---

[57] Spooner as well believes that Butler doesn't consider growth or development in the conscience or moral sense, but assumes that it is one and the same "in all men, whose judgment is not perverted by self-partiality [...] That was the unhistorical eighteenth century way of regarding the matter." (1901, p. 113)

time and space; and when he cites Epictetus as his reference for the term "moral approving and disapproving faculty" ((1), pp. 246-7), he is going back 16 centuries, whereas Duncan-Jones comes only two centuries after Butler! W. Lucas Collins is of like mind: "Butler was prepared to meet the real difficulty which lies upon the threshold of his doctrine,–that conscience is a shifting rule, varying with the various stages of civilization–with age, with country, and even with climate" (Collins quoted in Lefevre 1900, p. 402).

Today, however, there is an additional dilemma even assuming we intuitively know our duty. With God and religion out of the picture[58], we may wonder how to settle a conflict between what we believe will lead to the greatest overall happiness and what we believe to be in our own interest. This is not a question of motivation ("Why should I do my duty?"), but rather of a conflict of obligations. As Darwall states, without teleological metaphysics, there is no longer a guarantee that "an action's being something an agent *ought* to do, is *the same thing as* the action's furthering the agent's good" nor that there is "a harmony of individuals' goods", for that matter (1995, p. 3).

For this reason, Charles Larmore evaluates Butler's principal strength as having helped to lay the groundwork for the debate to follow between consequentialists and deontologists:

> Because Butler saw so distinctly the variety of moral reasons that we acknowledge, and because he made so clear how religion alone can secure their unity, he commands our attention in a way that no other moral philosopher in the eighteenth century (not even Kant, as we shall see) can do. He set the terms of the debate

[58] Which reminds me of an amusing inscription on the fence of an outdoor crêperie in Îles-de-la-Madeleine: "Nietzsche: 'God is dead.' God: 'Nietzsche is dead.' "

between deontology and consequentialism for the succeeding centuries. (1987, pp. 138-139)[59]

[59] See also Darwall 1995 at p. 4: "The modern conception of morality developed as a solution to the problem of conflicting interests, especially the problem of conflict among persons who cannot expect to share a common confession or religious discipline."

Conclusion

Butler's major contributions to ethical theory are his refutation of psychological egoism and his elaboration of the authority of moral conscience. Though our views of moral psychology have changed since the eighteenth century, his description of human nature still resembles our common sense evaluation of the workings of our mind in moral deliberation. Furthermore, in a similar fashion to Butler's contemporaries, we question the obligatory bind of conscience and whether we have good reason to follow it.

Philosophically speaking, this translates into scepticism towards moral standards which were written at a time when we feared the wrath of God and hellfire. With this doubt comes a renewed defence of the virtue of egoism and, at the same time, of its counterpart, altruism. Why all the debate? Are we not just as we are, to paraphrase Butler?

Self-made man is perpetually in quest of the purpose of his existence. Butler's knowledge that conscience has supreme authority needs to be justified, since the teleological assumptions which were valid in his day can no longer serve as proofs. Why then do we find his arguments so appealing?

Butler's theory has the advantage of discovering an inner or internal moral authority. The autonomous agent is not required to seek direction externally: the answer is found within. Does this, however, entail moral relativism? Not according to Butler's view of human nature, and he may not be that far off from the truth. Butler's answer is not: "Do what you want." It is "Let your conscience be your guide". This then raises the question of the source of conscience's dictates, and whether they are innate or learnt.

Another lesson from which we profit in Butler's naturalistic teleology is that benevolence is not the enemy of self-love we imagine it to be. This is one of the unfortunate results of our new religion, capitalism, in that materialism

translates itself into a desire for exclusive ownership of non-material goods. Butler reminds us that we can happily share the air, the light of the sun, and the blessings in life (peace, plenty, freedom, healthful seasons), without any loss to ourselves (Sermon XI, (13), p. 175 and Sermon XII, (7), p. 188). So jealous are we of our own happiness that we often miss out on occasions to improve it.

Unfortunately, Butler cannot escape the circularity of his argumentation because he has chosen a methodology which looks to the world and then states things as it finds them. He makes a very brave attempt to demonstrate that the authority of conscience is not arbitrary, but at the same time is aware that every "honest man" must judge for himself, and that at times the self-evident cannot be justified further.

Butler's confidence that duty and interest coincide is not purely theological. It is based on illustrations that the two are not only compatible, but enablers of one another. However, Butler seems to say in the "cool hour passage" that irreconcilable conflicts between the two can only be resolved by erring on the side of interest. Our first duty and care is to ourselves, since we are better equipped to help ourselves than any other person and vice versa.

To use Butler's metaphor, how do we know that a watch's purpose is to tell time? We know this because we designed the watch ourselves, based on our concept of time, which we constructed following what we found in the world, *i.e.* night, day, and the passage of time. How do we know that our life's purpose is virtue or the good? We know this because we are the designers of our own life and we have chosen virtue, based on its suitability to our nature: the fact that those under our care, by their very proximity, are our responsibility; and that we are under particular obligation to ourselves as well.

It is occasionally difficult to separate the normative from the descriptive in Butler's argumentation. He attempts to refute psychological egoism, and in the same breath, appears to recommend it. Butler's account of human nature is

normative, in that the universal moral standards which are written in our nature can be discovered by a thorough examination of it. Just as we cannot justify to ourselves an action against our own interest, Butler feels that he cannot recommend such an action as our duty.

Butler believes that we are well adapted to justice, veracity, regard to common good and prudence, for we condemn falsehood, violence and injustice, apart from their consequences, and distinguish between merit and ill desert. Vicious actions are thus disproportionate or unfit to our nature. Butler's naturalism derives moral duties from our commonplace approvals and disapprovals. What is implicit in his method is that eternal immutable moral standards exist and can be discovered and justified over time. He believes that, generally speaking, we are in moral agreement concerning our duties. This certitude is however diminished today, since we are exposed to a greater variety of mores, and our own are ever-evolving as well.

Butler's theory leaves us questioning the source, perhaps not of the *authority*, but of the *content* of the dictates of our conscience. If they do not originate in divine legislation, nor in societal or contractual norms, but in the "nature of things", what in nature is to serve as a norm? Moreover, to answer this, must we look to evolutionary biology, experimental psychology, philosophy of mind? Furthermore, is Butler's list of standards of virtue exhaustive? How are they to be applied to concrete situations? What do we do in a case of conflict of duties? These are all problems Butler does not attempt to resolve, due to his faith in the honest man's conscience. After all, he has left philosophers with a very rich inheritance indeed.

BIBLIOGRAPHY

A) Joseph Butler's Published Works

BUTLER, Joseph. *Works*:

The Analogy of Religion, Natural and Revealed, to the Constitution and Course of Nature.
To which are added, Two Brief Dissertations:
I. On Personal Identity.—II. On the Nature of Virtue.

Sermons:
Fifteen Sermons Upon Human Nature, or Man Considered as Moral Agent.
Six Sermons Preached Upon Public Occasions,
To which are added,
A Charge to the Clergy of the Diocese of Durham, 1751.
Correspondence between Dr. Butler and Dr. Clarke.

- ed. Samuel Halifax. London: Scott, Webster & Geary, 1838.
- ed. William Ewart Gladstone. 2 vols. Oxford: Clarendon Press, 1896.
- ed. John Henry Bernard. 2 vols. London: Macmillan & Co., Ltd., 1900.

BUTLER, Joseph. *Fifteen Sermons Preached at the Rolls Chapel and a Dissertation Upon the Nature of Virtue*, ed. The Very Rev. W. R. Matthews. London: G. Bell & Sons Ltd, 1967.

B) Books Consulted

1) On Butler's Ethics:

BAKER, Albert Edward. *Bishop Butler*. London: Society for Promoting Christian Knowledge, 1923.

CARLSSON, P. Allan. *Butler's Ethics*. The Hague: Mouton & Co., 1964.

COLLINS, W. Lucas. *Butler*. Edinburgh: William Blackwood and Sons, 1881.

COOK, Webster. *The Ethics of Bishop Butler and Immanuel Kant*, Philosophical Papers of the University of Michigan, Second Series, No. 4. Ann Arbor: Andrews & Company, Publishers, 1888.

CUNLIFFE, Christopher (ed.). *Joseph Butler's Moral and Religious Thought: Tercentenary Essays*. Oxford: Clarendon Press, 1992.

DUNCAN-JONES, Austin. *Butler's Moral Philosophy*. Harmondsworth: Penguin Books, 1952.

GLADSTONE, William Ewart. *Studies Subsidiary to the Works of Bishop Butler*. London: Henry Frowde, 1896.

JEFFNER, Anders. *Butler and Hume on Religion*, trans. K. Bradfield. Stockholm: Diakonistyrelsens Bokförlag, 1966.

MOSSNER, Ernest Campbell. *Bishop Butler and the Age of Reason*. New York: The Macmillan Company, 1936.

NORTON, William J., Jr. *Bishop Butler: Moralist and Divine*. New Brunswick: Rutgers University Press, 1940.

PENELHUM, Terence. *Butler*. London: Routledge & Kegan Paul, 1985.

SPOONER, W. A. *Bishop Butler: His Life and Writings*. Boston: Houghton, Mifflin & Co., 1901.

TAYLOR, William Edington. *The Ethical and Religious Theories of Bishop Butler*. Toronto: Bryant Press, Ltd., 1903.

WHYTE, Alexander. *Bishop Butler: an Appreciation with the Best Passages of his Writings*. Edinburg: Oliphant, Anderson & Ferrier, 1903.

2) <u>On Egoism/Altruism:</u>

NAGEL, Thomas. *The Possibility of Altruism*. Oxford: Clarendon Press, 1970.

ÖSTERBERG, Jan. *Self and Others: A Study of Ethical Egoism*. Dordrecht: Kluwer Academic Publishers, 1988.

RAND, Ayn. *The Virtue of Selfishness: A New Concept of Egoism, with additional articles by Nathaniel Branden*. New York: New American Library, 1964.

ROBERTS, Tom Aerwyn. *The Concept of Benevolence: Aspects of Eighteenth-Century Moral Philosophy*. London: The Macmillan Press Ltd., 1973.

SHAVER, Robert. *Rational egoism: A selective and critical history.* Cambridge: Cambridge University Press, 1999.

SOBER, Elliott and WILSON, David Sloan. *Unto Others: The Evolution and Psychology of Unselfish Behavior*. Cambridge: Harvard University Press, 1998.

3) <u>On Ethics:</u>

ARRINGTON, Robert L. *Western Ethics: An Historical Introduction.* Malden: Blackwell Publishers Inc., 1998.

AUDARD, Catherine (ed.). *Une anthologie historique et critique de l'utilitarisme. Les utilitaristes: leurs précurseurs et leurs critiques (1711-1903)*. Paris: Presses Universitaires de France, 1999.

BLACKBURN, Simon. *Ruling Passions: A Theory of Practical Reasoning*. Oxford: Clarendon Press, 1998.

BROAD, Charlie Dunbar. *Five Types of Ethical Theory*. London: Kegan Paul, 1930.

BROWN, Stuart (ed.). *British Philosophy and the Age of Enlightenment*, Routledge History of Philosophy, Vol. V. London: Routledge, 1996.

CANTO-SPERBER, Monique. *La philosophie morale britannique*. Paris : Presses Universitaires de France, 1994.

CANTO-SPERBER, Monique. *L'inquiétude morale et la vie humaine*. Paris : Presses universitaires de France, 2001.

DARWALL, Stephen. *The British moralists and the internal 'ought': 1640-1740.* Cambridge: Cambridge University Press, 1995.

ENGBERG-PEDERSEN, Troels. *The Stoic Theory of Oikeiosis: Moral Development and Social Interaction in Early Stoic Philosophy*. Denmark: Aarhus University Press, 1990.

EPICTETUS.
Discourses: Book I, trans. Robert F. Dobbin. Oxford: Clarendon Press, 1998.
Enchiridion, trans. George Long. Buffalo, N.Y.: Prometheus Books, 1991.

FLANAGAN, Owen. *Varieties of Moral Personality. Ethics and Political Realism*. Cambridge: Harvard University Press, 1991.

LARMORE, Charles. *Patterns of Moral Complexity*. Cambridge: Cambridge University Press, 1987.

MOORE, George Edward. *Principia Ethica*. Cambridge: Cambridge University Press, 1971.

PRICHARD, Harold Arthur. *Moral Obligation and Duty and Interest*. Oxford: Oxford University Press, 1968.

PRIOR, Arthur N. *Logic and the Basis of Ethics*. Oxford: Clarendon Press, 1949.

RAPHAEL, D. Daiches (ed.). *British Moralists, 1650-1800*. 2 vols. Oxford: Clarendon Press, 1969.

RASHDALL, Hastings. *The Theory of Good and Evil: A Treatise on Moral Philosophy*, 2nd ed. New York: Kraus Reprint Co., 1971.

SCHLICK, Moritz. *Problems of Ethics*, trans. David Rynin. New York: Dover Publications Inc., 1962.

SCHNEEWIND, Jerome B. *The invention of autonomy: A history of modern moral philosophy*. New York: Cambridge University Press, 1998.

SCHNEEWIND, Jerome B. (ed). *Moral Philosophy from Montaigne to Kant. An Anthology*. 2 vols. Cambridge: Cambridge University Press, 1990.

SCHNEEWIND, Jerome B. *Sidgwick's Ethics and Victorian Moral Philosophy*. Oxford: Clarendon Press, 1977.

SETH, James. *English Philosophers and Schools of Philosophy*. London: J.M. Dent & Sons, Ltd., 1912.

SIDGWICK, Henry.
Outlines of the History of Ethics for English Readers (6th ed.). London: Macmillan & Co. Ltd., 1962.
The Methods of Ethics (unabridged 1907 7th ed.). Indianapolis: Hackett Publishing Company, 1981.

STEPHEN, Leslie. *History of English Thought in the Eighteenth Century*. 2 vols. N.Y.: Harcourt, Brace & World, Inc., 1962.

WILLEY, Basil. *The Eighteenth Century Background: Studies on the idea of nature in the thought of the period.* London: Chatto, 1949.

C) Articles Consulted

1) On Butler's View of Conscience:

DARWALL, Stephen. "Conscience as Self-Authorizing in Butler's Ethics", in Cunliffe 1992, 209-241.

HEBBLETHWAITE, Brian. "Butler on Conscience and Virtue", in Cunliffe 1992, 197-207.

JAMES, Edward W. "Butler, Fanaticism and Conscience", *Philosophy*, 56 (1981), 517-532.

LEFEVRE, Albert. "Conscience and Obligation in Butler's Ethical System", *The Philosophical Review*, 9 (July, 1900), 395-410.

MARTIN, Michael W. "A Defense of the Rights of Conscience in Butler's Ethics", *Philosophy Research Archives*, 3, No. 1203 (1977), 90-101.

O'BRIEN, Wendell. "Butler and the Authority of Conscience", *History of Philosophy Quarterly*, Vol. 8, No. 1 (January, 1991), 43-57.

PRITCHARD, Michael S. "Conscience and Reason in Butler's Ethics", *Southwestern Journal of Philosophy*, 9 (Fall, 1978), 39-49.

RAPHAEL, D. Daiches. "Bishop Butler's View of Conscience", *Philosophy*, 24 (July, 1949), 219-238.

RORTY, Amélie Oksenberg. "Butler on Benevolence and Conscience", *Philosophy*, 53 (1978) 171-184.

SAMS, Henry W. "Reflection", *The Philosophical Review*, Vol. 52, No. 4, 400-408.

STURGEON, Nicholas L. "Nature and Conscience in Butler's Ethics", *The Philosophical Review*, Vol. 85, No. 3 (July, 1976), 316-356.

SZABADOS, Béla. "Butler on Corrupt Conscience", *Journal of the History of Philosophy*, 14 (1976), 462-469.

TOMKINS, Silvan Solomon. *Conscience, Self-Love, and Benevolence in the System of Bishop Butler*, Ph.D. Thesis in Philosophy. Philadelphia: University of Pennsylvania, 1934.

TOWNSEND, H. G. "The Synthetic Principle in Butler's Ethics", *International Journal of Ethics*, 37 (October, 1926), 81-87.

WORTHEN, J. F. "Joseph Butler's Case for Virtue: Conscience as a Power of Sight in a Darkened World", *Journal of Religious Ethics*, 23 (Fall, 1995), 239-261.

2) On Conscience:

DUNCAN-JONES, Austin. "The Notion of Conscience", *Philosophy*, 30 (April, 1955), 131-140.

FUSS, Peter. "Conscience", *Ethics*, 74 (January, 1964), 111-120.

LEITES, Edmund. "Conscience, Casuistry and Moral Decision: Some Historical Perspectives", *Journal of Chinese Philosophy*, 2 (December, 1974), 41-58.

THILLY, Frank. "Conscience", *The Philosophical Review*, 9 (January, 1900), 18-29.

3) On the Status of Benevolence in Butler's Theory:

FREY, Raymond Gillespie. "Butler on Self-Love and Benevolence", in Cunliffe 1992, 243-267.

McNAUGHTON, David. "Butler on Benevolence", in Cunliffe 1992, 269-291.

RIDDLE, Glenn K. "The Place of Benevolence in Butler's Ethics", *Philosophical Quarterly*, 9 (October, 1959), 356-362.

4) On Psychological and Ethical Egoism / The "Cool Hour Passage":

BROAD, Charlie Dunbar. "Self and Others" (1953), *Broad's Critical Essays in Moral Philosophy*, ed. David R. Cheney. London: George Allen & Unwin Ltd, 1971, 262-282.

BROAD, Charlie Dunbar. "Egoism as a Theory of Human Motives", *Ethics and the History of Philosophy: Selected Essays*. London: Routledge and Kegan Paul Limited, 1952, 218-231.

FEINBERG, Joel. "Psychological Egoism", *Reason and Responsibility: Readings in Some Basic Problems of Philosophy* (10th ed.), ed. Joel Feinberg and Russ Shafer-Landau. Belmont, CA: Wadsworth Publishing Company, 1999, 493-505.

HARRISON, Jonathan. "Self Interest and Duty", *Australasian Journal of Philosophy*, 31 (May, 1953), 22-29.

HENSON, Richard G. "Butler on Selfishness and Self-Love", *Philosophy and Phenomenological Research*, Vol. 49, No. 1 (September, 1988), 31-57.

JACKSON, Reginald. "Bishop Butler's Refutation of Psychological Hedonism", *Philosophy*, 18 (July, 1943), 114-139.

JOHNSON, Wayne G. "Psychological Egoism: *Noch Einmal*", *Journal of Philosophical Research*, Vol. 17 (1992), 239-264.

KLEINIG, John. "Butler in a Cool Hour", *Journal of the History of Philosophy*, 7 (1979), 399-411.

LEITES, Edmund. "A Problem in Joseph Butler's Ethics", *Southwestern Journal of Philosophy*, 6, No. 2 (Summer, 1975), 43-57.

PHILLIPS, David. "Butler and the Nature of Self-Interest", *Philosophy and Phenomenological Research*, Vol. 60, No. 2 (March, 2000), 421-438.

SCOTT-TAGGART, M. J. "Butler on Disinterested Actions", *The Philosophical Quarterly*, Vol. 18, No. 70 (January, 1968), 16-28.

SLOTE, Michael Anthony. "An Empirical Basis for Psychological Egoism", *The Journal of Philosophy*, Vol. 61, No. 18 (October, 1964), 530-537.

SOBER, Elliott. "Hedonism and Butler's Stone", *Ethics*, 103 (October, 1992), 97-103.

STEWART, Robert M. "Butler's Argument Against Psychological Hedonism", *Canadian Journal of Philosophy*, Vol. 22, No. 2 (June, 1992), 211-221.

WHITE, Alan R. "Conscience and Self-Love in Butler's Sermons", *Philosophy*, 27 (October, 1952), 329-344.

5) On Butler's View of Human Nature:

BRINTON, Alan. " 'Following Nature' in Butler's Sermons", *The Philosophical Quarterly*, Vol. 41, No. 164 (July, 1991), 325-332.

LEFEVRE, Albert. "The Significance of Butler's View of Human Nature", *The Philosophical Review*, Vol. VIII, No. 2 (March, 1899), 128-145.

MILLAR, Alan. "Butler on God and Human Nature", in Cunliffe 1992, 293-315.

MILLAR, Alan. "Reply to Brinton's '"Following Nature" in Butler's Sermons' ", *The Philosophical Quarterly*, Vol. 42, No. 169 (October, 1992), 486-491.

MILLAR, Alan. "Following Nature", *The Philosophical Quarterly*, Vol. 38, No. 151 (April, 1988), 165-185.

6) On Butler's Ethics:

CARRAU, L. "*La Philosophie de Butler*", *Revue Philosophique de la France et de l'étranger*, 21 (1886), "I.–*La Morale*", 144-158 and "*L'Analogie*", 265-280.

FUSS, Peter. "Sense and Reason in Butler's Ethics", *Dialogue*, 7, No. 2 (September, 1968), 180-193.

GRAVE, S. A. "The Foundation of Butler's Ethics", *The Australasian Journal of Philosophy*, Vol. 30, No. 2 (August, 1952), 73-89.

LEMOS, Ramon M. "Butlerian Meditations", *The Southern Journal of Philosophy*, Vol. 41, No. 2 (Summer, 2003), 199-221.

McPHERSON, Thomas H. "The Development of Bishop Butler's Ethics", Part I, *Philosophy*, Vol. 23, No. 87 (October, 1948), 317-331; Part II, *Philosophy*, Vol. 24, No. 88 (January, 1949), 3-22.

MONCK, W. H. S. "Butler's Ethical System", *Mind*, Vol. 3, No. 11 (July, 1878), 358-369.

NEWBERRY, Paul A. "Joseph Butler on Forgiveness: A Presupposed Theory of Emotion", *Journal of the History of Ideas*, Vol. 62, No. 2 (April, 2001), 233-244.

TAYLOR, A. E. "Some Features of Butler's Ethics", *Mind*, Vol. 35, No. 139 (July, 1926), 273-300.

7) On Ethics:

ANSCOMBE, Gertrude Elizabeth Margaret. "Modern Moral Philosophy", *Ethics* (1954), repr. in *Collected Papers*, III: *Ethics, Politics and Religion* (Oxford, 1981), 26-42.

BRINK, David O. "Sidgwick's Dualism of Practical Reason", *Australasian Journal of Philosophy*, Vol. 66, No. 3 (September, 1988), 289-307.

DAVIDSON, Rev. W. L. "Separation of Questions in Philosophy", *Mind*, Vol. 9, No. 36 (October, 1884), 548-563.

FALK, W. D. " 'Ought' and Motivation", *Proceedings of the Aristotelian Society*, 48 (1948), 111-138.

FRANKENA, William K. "Sidgwick and the Dualism of Practical Reason", *The Monist*, Vol. 58, No. 3 (1974), 449-467.

HUGHES, Gerard J. "Natural Law", *A New Dictionary of Christian Ethics*, ed. James F. Childress and John MacQuarrie, London: The Westminster Press, 1986, 412-414.

SCHNEEWIND, Jerome B. "The Use of Autonomy in Ethical Theory", in *Reconstructing individualism: Autonomy, individuality, and the self in Western thought*, ed. Thomas C. Heller, Morton Sosna, and David E. Wellbery. Stanford: Stanford University Press, 1986, 64-75.

SOSOË, Lukas. "Henry Sidgwick et le problème du dualisme de la raison pratique", *Cahiers de philosophie politique et juridique*, No. 18, 1990, 63-76.