A Thesis

entitled

The Dream of a Scientific Ethics

by

Jorge Piaia Mendonca Junior

Submitted to the Graduate Faculty as partial fulfillment of the requirements for the

Master of Arts Degree in Philosophy

_____

Dr. John Sarnecki, Committee Chair

_____

Dr. Madeline Muntersbjorn, Committee Member

_____

Dr. James Campbell, Committee Member

_____

Dr. Amanda C. Bryant-Friedrich, Dean
College of Graduate Studies

The University of Toledo

May, 2018

An Abstract of

The Dream of a Scientific Ethics

by

Jorge Piaia Mendonca Junior

Submitted to the Graduate Faculty as partial fulfillment of the
requirements for the Master of Arts Degree in Philosophy

The University of Toledo
May 2018

In the present thesis, I analyze the idea of using descriptive/indicative statements, rooted in natural sciences, as a mean to derive prescriptive/normative moral statements. In the first chapter, I examine different attempts to develop an account of ethics based on scientific knowledge. I try to explain *how* some philosophers and scientists, namely, August Comte, Moritz Schlick, and Sam Harris, argued for a science-based ethics, *why* they were inclined to do it, and *which problems* their projects came across. I argue that each of these attempts fails to reach its goal. In the second chapter, I look at the work of the biologist/political philosopher Peter Kropotkin, as a case study. I also explain the development of the concept of altruism in science, addressing the limitations of different concepts of altruism. In this second chapter, I try to demonstrate how scientists can be misled by scientific results in the pursuit of the dream of a scientific ethics.

Ao meu irmão, Gabriel, por me dar razões para comprar a briga contra a impassível mesquinharia que é engrenagem do mundo e por me fazer ver como a dicotomia egoísmo-altruísmo é redução grosseira do abismo sem penhascos que é o afeto.

# Contents

# Preface

The astonishing success of scientific knowledge, impacting every sector of life, reasonably encourages our trust in science. But this reasonable trust sometimes mutates into an unreasonable optimism. This mutation is not always clear and assumes different forms. In this thesis, I address an idea that I consider to be a consequence of such optimism: the idea that science, in addition to telling us how things *are*, can tell us how they *should* be. In other words, I address the idea of using descriptive/indicative statements, based on natural sciences, as a mean to derive prescriptive/normative moral statements.

The fundamental problem that will be addressed in this thesis can already be found in the 18th century, in the work of David Hume. In his *A Treatise of Human Nature* (1738), Hume comments that some philosophers claimed that "morality is susceptible of demonstration", and, even though "no one has ever been able to advance a single step in those demonstrations", these philosophers had "taken for granted that this science may be brought to an equal certainty with geometry and algebra" (Hume, 1738, p. 330). Hume argued against the possibility of deriving morality from pure reason[1], presenting different

1. This thesis will focus on the project of a scientific-based ethics, and not much on projects that want to derive ethics from pure reason, even though both projects often overlap significantly.

arguments against it. For the purposes of this thesis, one specific problem mentioned by Hume will be particularly important. This problem is known as the "*is-ought problem*".

> In every system of morality, which I have hitherto met with, I have always remark'd, that the authors proceed for some time in the ordinary way of reasoning… when of a sudden I am surpriz'd to find, that instead of the usual copulations of propositions, *is*, and *is not*, I meet with no proposition that is not connected with an *ought*, or an *ought not*. This change is imperceptible; but is, however, of the last consequence. For as this *ought* or *ought not*, expresses some new relation or affirmation, 'tis necessary that it should be observ'd and explain'd; and at the same time that a reason should be given, for what seems altogether inconceivable, how this new relation can be a deduction from others, which are entirely different from it (Hume, 1738, p. 334).

The attitude that Hume is reporting here is the attempt to deduce an "*ought*" from an "*is*". Using more a more precise terminology, it is the attempt to derive imperative/prescriptive statements from descriptive/indicative statements.

The British philosopher G. E. Moore (1903) introduced a similar problem in reasoning, which he called the "naturalistic fallacy". While Hume focused on the problem of deriving ought from *reason*, Moore focused on the problem of deriving ought from *nature*. In short, Moore's naturalistic fallacy consists in inferring a positive moral value from what is considered to be natural. In Moore's terms, the naturalistic fallacy was "the supposition that good can be defined by reference to a natural object" (1903, p. 39). For Moore, anyone who argues that "because something is natural it should be acceptable or have a moral value due to its status of being natural" was committing the naturalistic fallacy. To assume that we can derive normative statements from what is natural can be considered a *subset* of the more general problem of deriving normative statements from descriptive statements.

Years later, in the context of Philosophy of Science, Henri Poincaré presented the same "is-ought problem" focusing more on the linguistic aspect of the problem. He argues

that "[t]here cannot be a scientific morality", since "[i]f the premises of a syllogism are both in the indicative, the conclusion will also be in the indicative. For the conclusion to have been stated in the imperative, at least one of the premises must itself have been in the imperative" (Poincaré, 1913, p. 105). As Hare (1952, p. 5) explains, "[a]n indicative sentence is used for telling someone that something is the case; an imperative is not—it is used for telling someone to make something the case"[2].

In the context of science, or, more broadly, of natural philosophy, we assume that sciences are giving us *descriptions* of nature. Thus, the derivation of *normative* statements from the scientific body of knowledge is clearly problematic. Nevertheless, some philosophers believed that they could overcome this problem. The first chapter of this thesis will present how some of these philosophers—Auguste Comte, in the 19th century, Moritz Schlick, in the 20th century, and Sam Harris, in the 21st century—have argued for a scientific ethics. More than simply presenting the arguments of these authors, a broader understanding of the context in which such authors defended these ethical projects will be an important part of this thesis and will help the understanding of why the idea of a scientific ethics was (and still is) so seductive.

After addressing, in the first chapter, the *philosophical* development of the idea of a scientific-based ethics and different arguments for such an ethical project, the second chapter will focus on the *scientific* aspects of scientific-based ethics. The attempt to use scientific knowledge to justify morality isn't just a problem in moral philosophy, it can also be a problem for scientific research. Thus, if the first chapter will focus on the *philosophical*

2. Imperative statements aren't the only form of normative statements. An evaluation of something, like "this painting is good" is a normative statement, too. These differences will be addressed in the first chapter.

development of the idea of a scientific-based ethics, addressing examples of philosophers and their arguments for such ethical project, the second chapter will explore how moral values can have an influence on the scientific research and which mechanisms can be used in science in order to avoid such influence.

It is a fair assumption to consider that regardless of the methodological differences between different disciplines of scientific knowledge, one thing that we want to be able to say about a proposition is this: if it is a "scientific" proposition, then it is not discussing matters of *choice*. We expect from good scientific methods, from psychology to physics, that such methods will guarantee that what we are claiming scientifically will not be influenced by personal preferences. In short, we don't want scientific propositions to be influenced by the scientist's arbitrary *values*.

Notwithstanding the requirement above, we are obligated to recognize that there is always a plurality of overlapping *values* underlying scientific practice. In fact, even our rule that science should be free from values can be considered to be a value itself. As Lacey puts it, a value "to be expressed in scientific practices and embodied in scientific institutions, a value embedded in the objectives of science itself" (2005, p. 18). The fantasy of science as the pure use of reason and the myth of the scientist's disinterested eye looking to "reality as it is" is obsolete.

At this point, it seems that we have a contradiction: on the one hand, science is inevitably subjected to values, and, on the other hand, the belief that science "is about facts" instead of being about values seems to be a fundamental assumption. In this scenario, scientists would be engaged in the weird activity of producing propositions that are necessarily influenced by values and that, at the same time, should not be influenced by

values. A straightforward approach to avoid this apparent contradiction is to propose the distinction between "cognitive" and "non-cognitive" values[3]. I'll explain.

From the hypothesis that science *cannot* be "value free", one cannot infer that *any* sort of value should be allowed to influence scientific research. Values that are beneficial or necessary for scientific practice should be identified. There must be a *distinction* between the values we want to influence science and the values we don't. The distinction between "cognitive" and "non-cognitive" values is one way to make this distinction. Values such as simplicity, fruitfulness, predictability, coherence, etc., are intrinsic parts of science and should be pursued by the scientist: these are the ones we call "cognitive values" (see Lacey, 2005, p. 45-51). Discussions about which of these values should be more important and strategies for better pursuing them are fundamental.

The presence of what is called "cognitive value" is not a problem at all: they are necessary for science and we cannot imagine science without them. What we really want when we say that science should be value-free, is the rejection of a particular kind of values, namely, the values called "'non-cognitive'" (Lacey, 2005, p. 15). The set of arbitrary values, which we usually identify as "personal preferences", is what can influence negatively science. Even though non-cognitive values are not necessarily moral—they can be aesthetical, for example—, in this thesis I'll focus on the moral ones.

In the second chapter of this thesis, I'll explore some possible influences of non-cognitive values in scientific research. The first case study will be the work of Peter Kropotkin, at the beginning of the 20th century. In this first section, I will explore some

3. This distinction is far from being perfect. However, the complexity of the relation between science and values will not be addressed here in detail. For the purposes of this thesis, this broad distinction will be enough. For a deeper view on this, see Lacey (2005).

possible flaws of early evolutionary biology and how such flaws could allow non-cognitive values to influence science. The second case study will focus on the term "altruism" in contemporary biology and psychology, exploring the benefits and limitations of the recent scientific terminology.

# Chapter 1

# The Dream of a Scientific Ethics

> During the latter half of the nineteenth century, people often dreamed of formulating a scientific ethics. We were not content to sing the praises of the educational virtue of science, the advantages that the human soul derives for its own improvement from looking truth in the eye. We relied on science to place moral truths beyond all contestation as it has done for the theorems of mathematics and the laws stated by the physicists (Poincaré, 1913, p. 102).

This chapter will address some philosophers who share what Poincaré called the "dream of a scientific ethics". In short, the dream of a scientific ethics is the idea of deriving normative moral statements from scientific knowledge. Instead of the broad problem of deriving morality from reason, this chapter will address some cases where philosophers tried to derive morality specifically from scientific-based knowledge. Each of the following three sections will consist of an investigation of a different versions of this "dream" in a different context.

The method used here will *not* be the purely epistemological approach of analyzing arguments by themselves—the understanding of a philosophy goes beyond the simple analysis of its arguments. If we consider a philosophy as a product of a given socio-cultural context, we need to understand the influence of this context on this philosophy in order to understand this philosophy. Here, I argue that we can identify a pattern, namely, the "dream

of a scientific ethics", which repeats itself in different philosophies. I'll try to show *how* philosophers defended such ideal, *why* they did it, and, finally, explain why their projects were condemned to fail. The choice to call it by the term "dream", inspired by Poincaré, is a way to emphasize both the greatness of such project and its delusional character.

It will be important, before anything, to stress two distinctions that will be helpful in order to avoid a too simplistic critique of the dream of a scientific ethics. After these distinctions, some bad arguments against the scientific ethics will be discarded, leaving room for the important problems of such a project to be discussed later on.

The first distinction regards the meaning of imperative statements. In his book *The Language of Morals* (1952), R. M. Hare makes an important distinction between two ways of interpreting imperative statements, i.e., two ways of *describing* what imperative statements actually mean. "The first does this [description of imperatives] by representing them as expressing statements about the mind of the speaker. Just as it has been held that 'A is right' means 'I approve of A'" (Hare, 1952, p. 5). And "[t]he second attempt... may be summarized… by the statement that 'Shut the door' means the same as 'Either you are going to shut the door, or X will happen', where X is understood to be something bad for the person addressed" (Hare, 1952, p. 7). Therefore, the usage of imperatives isn't necessarily the straightforward imposition of one's will over someone else's: it can mean something more complex, involving conditionals. With this in mind, let's consider another distinction.

The second distinction regards ways of deriving norms from descriptions. One of the possible ways to make people accept a set of *norms* deliberately (without forcing or misleading them) is to present them with a *description* of the world from which such norms

would follow. The more our description is justified, the more a norm that follows from it will be justified as well. The idea of a God and the idea of a soul are good examples. We can describe reality as having souls and a God, and, then, considering such description as our background, we can explain that some norms (of not committing suicide, for example) simply follow from reality as it is. Therefore, in this example, when we affirm the rule "you should not commit suicide", we are just claiming that this is simply how reality is—reality is such that suicide is wrong.

But the example above is *not* the method used by the authors who argued for the scientific ethics. The basic structure used by the proponents of the scientific ethics is something like this: (1) they *describe* the existence of some motivations (m) in the individual (e.g., seeking pleasure), then (2) they *describe* some behaviors (b) as being the ones that would better achieve what these motivations (m) aim to, and then, finally, (3) they *prescribe* the behaviors (b) as a moral norm (n). Then, they say: "considering the conjunction between particular facts discovered by science about behaviors (b) and motivations (m), the norm (n) follows". The philosopher uses, then, imperative statements to prescribe the behaviors (b).

As we can see, considering Hare's distinctions between imperatives, the prescription above is not straightforward imposition of one's will, but a complex conditional statement. A norm like "love your neighbor", which could be stated as "it is right to love your neighbor", would mean "if you want the best for yourself, then you should love your neighbor". We can also say it in another way: "either you are going to love your neighbor, or X will happen", where X is something that is claimed to be objectively bad for yourself.

The main advantage of the second way of deriving norms (using conditionals) is that the proponent of a scientific ethics can claim that scientific ethics is not a restriction on one's freedom[4]. In the first way of deriving norms, even if we recognize that a given norm is the will of God, it is still an order reducing the individual's freedom. But the proponents of a scientific ethics argue that the norms of the scientific ethics are just the action that *we would do if* we knew some facts about ourselves—they claim that these "scientific-based norms" are just stating what we *actually* want. I'll try to explain why their reasoning is flawed.

After the introduction to the project of a scientific ethics stated above, we can address some historical cases in which we can identify its occurrence. It is almost impossible to have a proper understanding of the dream of a scientific ethics and the underlying motivations for such project without understanding first the main ideas of positivism. Considering this, the first case study will be Auguste Comte's Philosophy. The pressure that comes from the rejection of metaphysics and the evaluation of science can be taken as the major factor for the attempt to underpin morality in nature. Comte's case will also be especially interesting, here, due to the simplicity—perhaps naivete—of Comte's defense of a scientific-based ethics.

4. It is important to mention that this way of deriving norms can be used by religious-based arguments as well. For example, consider that the motivation to avoid suffering is assumed by the philosopher. Now, consider that the reality he is describing is such that we do have souls, and suicide makes such souls burn in hell for eternity. Since people don't want to suffer and suicide would make them burn in hell, then the norm "do not commit suicide" is simply what we would do *if we knew the world as it is*, i.e., if we properly knew the causal relations of reality.

### 1.1. Comte, Positivism, and Altruism

In a passage explaining the general goals of his Positivist Philosophy, Comte says that "[t]he primary object, then, of Positivism is twofold: to generalize our scientific conceptions, and to systematize the art of social life" (1848, p. 3). The generalization of our scientific conceptions, explaining human knowledge as an organized *system*, is, perhaps, the most famous aspect of Comte's philosophy. But Comte was also interested in "the art of social life", and the scope of his philosophy includes problems regarding morality and sociology.

A systematic presentation of history, explaining the patterns underlying historical processes, was one of the main goals of Comte. His Law of the Three Stages (Comte, 1842), or sometimes Laws of Human Development, was crucial to his Positive Philosophy (see Sharma, 1996, p. 49-51). With this law, Comte wanted to explain the rule of the unfolding of human society. For him, the three stages represented the "progressive course of human mind" (Comte, 1842, p. 27), and every form of human knowledge passed—or will pass—through all three stages.

> From the study of the development of human intelligence, in all directions, and through all times, the discovery arises of a great fundamental law [Law of Three Stages], to which it is necessarily subject, find which has a solid foundation of proof, both in the acts of our organization and in our historical experience (Comte, 1842, p. 27).

In the first of these stages, which Comte called the Theological Stage, society believed in gods and spirits underlying nature. The first stage was the stage in which "free play is given to spontaneous fictions admitting of no proof" (Comte, 1848, p. 34-35). Comte considered this stage as a primitive level of social and intellectual development, in

which a “lack of logical and orderly thinking” (Sharma, 1996, p. 49) allowed supernatural explanations to be accepted.

In the second stage, called the Metaphysical Stage, the supernatural elements are substituted with rational concepts. In the Metaphysical Stage, vague notions from the Theological Stage, such as “god”, will be represented using abstract concepts. It is a stage “characterized by the prevalence of personified abstractions or entities” (Comte, 1848, p. 35), and even though the questions of the second stage remain the same as the first stage, now they are answered in terms of abstract ideas, e.g., god as the first cause.

Finally, the Positive Stage, or the Scientific Stage, excludes the supernatural: it is the stage “based upon an exact view of the real facts of the case” (Comte, 1848, p. 35). Here, the metaphysical speculations and the revelations from religion are simply irrelevant. For Comte, in the third stage humanity reaches a level of awareness that results in the acceptance of the knowledge based on observation, experiment, and comparison as the proper way to know reality.

The general Law of Three Stages was not only applied at the level of societies: it was applied to the individual level as well. Comte claims that, in an analogous way to the development of history, “each of us is aware, if he looks back upon his own history that he was a theologian in his childhood, a metaphysician in his youth, and a natural philosopher in his manhood” (Comte, 1842, p. 29).

Comte stated that the purpose of his Positive Philosophy was “to supersede Theology in the spiritual direction of the human race” (Comte, 1848, p. 60) and that it was “the only rational means of exhibiting the local laws of the human mind” (Comte, 1842, p. 35). Considering Comte’s understanding of his Positivism, it is not hard to see how

attractive the idea of constructing a moral doctrine free from theology and metaphysics was for him. The Law of the Three Stages wasn't just a way of explaining the past. Comte's Positivism wasn't just a speculative doctrine: he wanted active practical use for it. The distinction between speculative and practical philosophy was already explained by Hume (1738, p. 325):

> Philosophy is commonly divided into speculative and practical; and as morality is always comprehended under the later division, 'tis suppos'd to influence our passions and actions, and to go beyond the calm and indolent judgments of the understanding.

For Comte, the "'great problem of human life' was to organize society so that it functions well without being disrupted by selfish interests" (Wilson, 2015, p. 142-143). Applying the Law of the Three Stages to this problem, Comte states that "[t]here are three successive states of morality answering to the three principal stages of human life; the personal, the domestic, and the social stage" (Comte, 1848, p. 100). He claims that selfishness was part of an early stage of development, and the progress would help to combat this selfish instinct.

But while Theologians and Metaphysicians attempted to justify their normative claims through supernatural explanations or by purely abstract constructs, such strategy would not be accepted by the philosopher of the Positive Stage[5]. Comte's philosophy was

5. In order to avoid comparing Comte only with the Moral Philosophers, an interesting analogy can be made between Comte and the American writer R. W. Emerson. In opposition to the Calvinist doctrine, which stated that humans were inherently despicable sinners, Emerson proposed, in his lectures and poems, that humans have the real essence of religion inside themselves, and, from the contact with this essence, they would be able to extract their own morality, without the need of religion (see Emerson, 1940, p. 471-500). For Emerson, the rigidness of the tradition only obliterated the good essence of human beings. This analogy is instructive to show that, while both authors opposed themselves to theology, Emerson didn't oppose himself to metaphysics. Using Comte's classification, Emerson still was in the second stage, while Comte was in the third one.

the manifestation of the Third Stage, therefore, an arbitrary morality, grounded in metaphysical ideas, wasn't a valid option. Thus, considering that for him the human being was something to be understood scientifically as any other object in nature, it was only in natural facts that the Positivist Comte could expect to underpin his morality. And it was the idea of *altruism*, as something innate and natural in human psychology, that Comte used for his project (see Wilson, 2015).

> Standing in direct connection with the fundamental principle of Positivist synthesis, the doctrine of innate altruism alone enables us to establish a systematic morality, which… may take the presidency, subjectively, of the encyclopedic hierarchy (Comte, 1877, p. 18).

The term "altruism" was coined by Auguste Comte in the middle of the 19th century. While egoism, or selfishness, is an intention or desire to benefit oneself, ignoring possible harmful consequences to the others, normally associated with a negative moral value[6], the word altruism was supposed to mean the exact antonym of egoism (Wilson, 2015, p. 4). Altruism was defined by Comte as "'living for others' (*vivre pour autrui*)" (Campbell, 2006, p. 359), or, more precisely, as "an unselfish regard for the welfare of others"[7] (Rhode, 2005, p. 56).

---

6. A famous example of argument for the positive moral value of egoism can be find in Ayn Rand's work (see Rand, 1964). For a critique of Rand's conception of altruism and a comparison to Comte's conception, see Campbell (2006).

7. Even though the term altruism was coined by Comte, the idea of acting for the other's sake can be found before him. In the tradition of Confucianism, we already can see a debate about what was called "universal love", i.e., the moral concern for others independently from the degree of relatedness. As Dubs (1951) shows, Confucius defended the idea that we should love our relatives, while his opponent, Mozi, defended the idea of universal love. The debate about universal love in the Post-Confucius tradition is closer to the discussion presented here than it may sound. We can find even the discussion about what is natural or not. As Dubs (1951) says, Mencius (372-298), a representative of the orthodox Confucianism, while arguing against Mozi's idea of universal love, stated that "equal love for all is unnatural. People naturally love their own parents more than those of others.

For Comte, a genuine altruistic attitude toward others was the mature posture, in opposition to the childish fear of divine punishment as a motor for a good behavior. The Catholic doctrine, as a part of a previous stage, was supposed to be substituted with the Positivist doctrine: "[w]herever evolution, individual or collective, follows its normal course, love first leads us to faith…. But when it becomes *systematic*, then the belief is constructed to regulate love" (Comte, 1891, p. 59). Comte criticized Catholicism not only for its metaphysical and theological justification, but for its moral consequences as well. With the concept of altruism, Comte opposed himself to what he considered to be "a direct contradiction of Catholic doctrine, which taught that human nature was entirely sinful and that love of others was only available through divine grace" (Wilson, 2015, p. 90).

Comte believed that the natural status of altruism would be proved through the biological research on animal behavior and what was the "young science of the brain", named "phrenology"[8] (Wilson, 2015, p. 90). For him, "that what theologians described as the struggle between the law of the flesh and the law of God", just like what metaphysicians described in their abstractions, "could be replaced by the scientific distinction between egoistic instincts located in the posterior part of the brain and altruistic instincts located in the anterior part of the brain" (Wilson, 2015, p. 90). Comte was so amazed by the idea of

---

Everyone has certain special duties to his own parents" (1951, p. 51). In Comte's view, Confucius was walking about the second stage (domestic stage), while Mozi was talking about the third stage (social stage) (see Comte, 1848, p. 100-102).

8. Phrenology was the study the mind through measurements of the skull. It was the combination of the right assumption that each part of the brain has a specific function and the wrong assumption that the format of the skull indicates the development of specific mental faculties. It was a popular practice in the 19th century, but nowadays it is considered to be a pseudoscience.

innate altruism that he even said that "[t]he innateness of the benevolent instincts and the earth's motion are the most important results of modern science" (Comte, 1877, p. 18).

If humans are altruistic by their very nature, then religion, God, and the whole metaphysical vocabulary can be dismissed. Positivism would articulate scientific knowledge with our psychological inclinations, and both, in this doctrine, would be in harmony. As Comte explains:

> The habitual predominance of altruism over egoism, in which lies the great problem for man, is in Positivism the direct result of a constant harmony between our best inclinations and all our labours, theoretical as well as practical (Comte, 1891, p. 47).

The chief characteristic of Positivism, in Comte's words, "consists in finally condensing, in one and the same formula, *the law of duty and the law of happiness*" (Comte, 1891, p. 215). To underpin morality in what makes humans happy, considering them as biological creatures and not as souls, is a common idea in the different projects of scientific ethics. The idea underlying Comte's moral project was the assumption that if we recognize our natural inclinations and these inclinations are good, then all we need is to, somehow, express these inclinations.

At this point, Comte had to explain how his doctrine would overcome the problem of *motivation*, i.e., the fact that the motivation to act morally doesn't follow necessarily from the *recognition* of the morality of an action. In the history of philosophy, we can divide the answers to this problem in two ways, namely, internalism and externalism.

> Internalism is the view that the presence of a motivation for acting morally is guaranteed by the truth of ethical propositions themselves…. Externalism holds, on the other hand, that the necessary motivation is not supplied by ethical principles and judgments themselves, and that an additional psychological sanction is required to motivate our compliance (Nagel, 1978, p. 7).

Since altruism is not an action that Comte claims to follow from pure reason, he can be identified as an externalist. Thus, even though altruism, if assumed as a natural fact, could be interpreted as a moral inclination in ourselves, the discovery of a natural altruistic impulse, by itself, wouldn't be enough to make people act altruistically. Comte believed that the recognition of such an altruistic nature would justify the normativity of his doctrine, and that the motivational problem could be solved through his Religion of Humanity. The center of the Religion of Humanity was "the feelings of venerative, identificatory and devotional love towards Humanity" (Wernick, 2001, p. 4). The "religious" activity of his Religion, which didn't propose any metaphysical entity, would have a specific function: "strengthen the altruistic impulses seen as vital for the correct orientation of thinking and acting" (Wernick, 2001, p. 4).

Comte is right when he says that an altruistic disposition towards others could lead to moral behaviors. He is also right to consider that a scientific-informed philosophy could explain the ways in which one can better express this altruistic drive. However, these properties of altruistic instincts don't justify what Comte tries to do with them—and here we can criticize Comte's project.

The problem of Comte's dream of a scientific ethics is simple: the existence of an altruistic instinct in human beings doesn't imply a positive moral evaluation of such instinct—this inference would clearly be a *naturalistic fallacy*. Even if it were the case, for example, that empirical data show that altruism is necessary for the existence of human beings, all that we would be able to infer from this are statements of the form "either we act altruistically or we will die". But the moral judgment which states that "to die" is "*bad*" is something that requires independent moral reasoning and is not logically implied in the

empirical data. To sum up: the evaluation of altruism as morally good doesn't follow from the observation of an altruistic instinct, i.e., there are not any logical implications between the scientific descriptions of altruism and the normative prescriptions of altruism.

### 1.2. Schlick and Ethics in The Vienna Circle

If some of Comte's ideas, such as the Religion of Humanity, didn't endure[9], some of them strongly influenced the philosophy of the 20th century. The core of the "Positivist spirit", with the rejection of Metaphysics and the evaluation of science as our primary form of knowledge, gained new momentum in Logical Positivism[10]. Certainly, ideas like a religion based in science wouldn't be popular for the Logical Positivists, but the "dream of a scientific ethics" was still present. In this section, I will address a new version of the "dream", this time in the context of the Vienna Circle, with the work of Moritz Schlick.

If one wants to list the characteristics shared by Comte's Positivism and the Logical Positivism of the Vienna Circle, the radical rejection of metaphysics[11] would be foremost. This rejection will be crucial for the scientific ethics, since it can be identified as the main reason that motivated these authors to search for the grounds for morality in nature.

9. "In practical terms, Comte's founding religious project was a complete, even preposterous, failure" (Wernick, 2001, p. 5).

10. "[T]he epithet 'logical' was added because they wished to annex the discoveries of modern logic; they believed, in particular, that the logical symbolism which had been developed by Frege, Peano and Russell would be serviceable to them" (Ayer, 1966, p. 10).

11. Differently from Comte's perspective, for the members of the Vienna Circle, there wasn't a significant difference between metaphysical and theological claims. Their rejection of Metaphysics, then, includes the rejection of Theology.

However, as Schlick (1948. p. 83) explains, "[i]f one wishes to characterize every view which denies the possibility of metaphysics as positivistic this is quite unobjectionable", but he adds that "this holds, of course, only under the presupposition of a special definition of 'metaphysics'". If on the one hand Comte's Positivism and the Vienna Circle shared the goal of rejecting metaphysics, on the other hand the Viennese definition of metaphysics and their reasons for such rejection differs from Comte's. Before presenting Schlick's argument for a scientific ethics, it will be important to explain what was, for him and for the rest of the Vienna Circle, the problem with metaphysics.

Following a tradition started with Kant, in the 18th century, the Logical Positivists believed that our knowledge of the world, or, in other words, the scope of what can be said, is restricted to the realm of "appearances", which these Positivists called "the given"[12]. As Hempel (1950, p. 108) states, "[t]he fundamental tenet of modern empiricism is the view that all non-analytic knowledge is based on experience". For the Logical Positivists, it was a condition of possibility for any *meaningful statement* to be either (1) derived from pure logic or (2) derived from empirical verification. Even if a statement is not verifiable at a given time, it should be "verifiable in principle" (see Hempel, 1966, p. 220-222). As Schlick stated, "the philosopher as well as the scientist must always remain within the given", and "to go beyond it, as the metaphysician attempts, is impossible or senseless" (Schlick, 1948. p. 83).

12. "The term 'the given' itself is a cause of grave misunderstandings…. the given is for him [the Positivist] but a word for what is most simple and no longer questionable" (Schlick, 1948. p. 84).

Thus, the Positivist's skepticism towards metaphysics was not a consequence of some sort of "metaphysical knowledge of the impossibility of metaphysics"[13], but rather the consequence of a simple recognition: if metaphysics was the attempt to reach the Absolute Truth, the Absolute Reality, or, in short, anything beyond the realm of "appearances", then metaphysical statements should not be considered as meaningful statements in the first place. Even though the vagueness and the presuppositions underlying every metaphysical claim were a problem, the Logical Positivists didn't reject metaphysical statements for this reason, but because it was impossible to determine the truth value of these statements, i.e., they were "meaningless sequences of words" (Schlick, 1930, p. 56).

Since metaphysical questions are not genuine questions, to get rid of them, including all the perennial moral questions from traditional philosophy, was not only possible but highly desirable. At the same time, the complete exclusion of the ethical questions from philosophy wasn't accepted by every member of the Vienna Circle. Schlick disagreed that ethics needed to be excluded from the set of meaningful topics. He argued that "[i]f there are ethical questions which have meaning, and are therefore capable of being answered, then ethics is a science" (1939, p. 1). The rejection of the old metaphysics, then, would not lead to a rejection of ethics, but only change the way in which we address its problems.

13. Some hasty critics accuse Logical Positivism of incurring in contradiction, arguing that the Vienna Circle's negation of metaphysics is already a metaphysical argument, thus, self-refuting itself. The members of the Circle were aware of this evident contradiction, but, since philosophy was conceived by them as an *activity* and not a system, such contradiction doesn't apply to their philosophy.

In *Problems of Ethics* (1939), Schlick presents his ideas regarding ethics. For Schlick, there were "no questions which are in principle unanswerable, no problems which are in principle insoluble" (Schlick, 1930, p. 56). And, at the same time that "Schlick agreed with Carnap, Ayer and the Vienna Circle that traditional ethics and value philosophy belong to metaphysics (have no empirical meaning)", he believed that ethics could be addressed in a different way, making it a meaningful area of knowledge: "the problem of ethics was not solved for him by a mere negative condemnation" (Leinfellner, 1985, p. 347).

For Schlick, if "the expression 'moral good' makes good sense", then "must we be able to discover it in a way analogous to that by which one discovers the meaning of the word 'life' or 'light'" (Schlick, 1939, p. 5). In other words, if we can talk about the moral good, then it is a natural phenomenon, and, if the moral good is a natural phenomenon, then its study must be scientific. For him, thus, instead of being a branch of philosophy, ethics should be considered a branch of science.

Schlick's ideas on ethics were mostly accepted by the other members of the Vienna Circle. As Ayer (1966, p. 22) states, "[t]he Vienna Circle as a whole was not very greatly interested in ethics; but it did not dispute Schlick's view that if ethical statements were to be brought into the scientific fold, they must be handled in the way that he proposed". However, there was still some skepticism about it: "[t]he only question was whether they [ethical statements] belonged within the [scientific] fold, whether they were statements of fact at all" (Ayer, 1966, p. 22).

As Leinfellner (1985, p. 318), in an analysis of Schlick's project in *Problems of Ethics*, explains, "[t]he central task of ethics according to Schlick, is to explain average

ethical behavior (ethical acts and decisions) by psychological and social law-like rules, i.e., rules for decision making". In *Problems of Ethics*, Schlick's "main argument is that human beings have solved this kind of conflict [ethical conflicts] since society began and long before ethics had been founded in Egypt as a doctrine of just actions and decisions" (Leinfellner, 1985, p. 320). Thus, assuming the *naturalness* of our capacity to solve conflict, Schlick infers that ethics, as it has been studied, is just "a reconstruction and rationalization of those solutions of social conflicts", and those solutions are "a decision procedure or mechanism deeply rooted in our psyche (i.e., in our imagination, or imagined representations in the conscious mind of the individual)" (Leinfellner, 1985, p. 320).

Perhaps a good way of explaining Schlick's ethics would be by comparing it with Utilitarianism, stressing some differences between both philosophies. Indeed, Schlick's model of ethics resembles Utilitarianism in many aspects: both doctrines agree, for example, in representing feelings by values (see Leinfellner, 1985, p. 327) and Schlick believed that "the ethical behavior of human beings is governed by positive and negative pleasure" (Leinfellner, 1985, p. 317). However, Schlick criticizes the Utilitarian formula "greatest happiness of the greatest number of human beings" saying that "the results of every act are simply incalculable, for they stretch on into time indefinitely; and even the resultant events of the near future cannot be predicted" (Schlick, 1939, p. 88). Accepting the resemblances between his project and classical Utilitarianism, Schlick makes a distinction:

> The formulation of our thesis is perhaps not unessentially different from that which it received in the classical systems of Utilitarianism. These systems say (at least according to their sense): "The good is what brings the greatest possible happiness to society." We express it more carefully: "In human society, that is *called* good which is *believed* to bring the greatest happiness" (Schlick, 1939, p. 87)

Another difference is pointed by Leinfellner (1985, p. 325):

> The real difference between maximizing utility and maximizing pleasure or satisfaction in the case of Schlick is that the former is an external maximization—very often regarded as the maximizing of greed—whereas Schlick's maximizing is the increase of an inner satisfaction and pleasure.

At this point, one could read Schlick's ethics as a study of the mechanisms involved in decision making, i.e., as a simply descriptive study. But if on the one hand Schlick's ethics seems to be just a psychological inquiry on the causes that make humans decide and solve conflicts, on the other hand, "[h]e is searching for those causes of ethical evaluations from which he will derive our ethical behavior" (Leinfellner, 1985, p. 327). Merely explaining human behavior is not what Schlick seems to have in mind, since "[f]or him decision theory and ethics… have to give an answer to the question, 'What shall we do?'" (Leinfellner, 1985, p. 318) and not only to the question "how do humans usually do what they do?".

Considering Schlick's context, where logic and language were objects of intense study, naturally he was aware of the "is-ought problem". He mentions the problem saying that "[i]n modern philosophy since Kant, the idea repeatedly appears that ethics as a normative science is something completely different from the 'factual sciences'", but he argues that "this manner of opposing normative and factual sciences is fundamentally false" (1939, p. 17). Schlick's argument—which would justify science's normative power—is to consider a different form of normativity: "in ethics we must be able to give the exact conditions under which the word 'good' is applied, even though its fundamental concept be indefinable" (Schlick, 1939, p. 9). Once the scientists discover the exact conditions of the good, they would be able to say that "[a] mode of action must have such

and such properties in order to be called 'good' (or 'evil')", and that "[s]uch a rule can also be called a 'norm.'" (Schlick, 1939, p. 14).

> The lawgiver who sanctions the moral commands is human society, which is furnished with the necessary power to command. Thus we may rightly say that morality makes demands on men, that they ought to behave in certain ways; because we use the word "ought" here in exactly the determined empirical sense (Schlick, 1939, p. 111).

Therefore, "if … the philosopher answers the question 'What is good?' by an exhibition of norms, this means only that he tells us what 'good' actually means; he can never tell us what good must or should mean" (Schlick, 1939, p. 18), or, in other words, "[e]thics must simply recognize this [the good] as a fact of human nature" (Schlick, 1939, p. 17). In Schlick's model, one would be able to determine the morally good behaviors scientifically, justified by the fact that these behaviors are, in fact, the ones that we humans really want "in the depths of our soul", even though we just don't know it yet[14]. The following passage makes this clear.

> We just said that there could be no real opposition between the meaning of the word "good" that is actually accepted in life, and the meaning found by the philosopher. An apparent difference can of course occur, for language and thought are very imperfect in daily life. Often the speaker and valuer is himself not clear as to what he expresses, and often his valuations rest on a false interpretation of the facts, and would at once change with a correction of the mistake. The philosopher would have the task of discovering such errors and faulty expressions, and would have to recognize the *true norms* [emphasis added] that lie at the root of moral judgments, and place them in opposition to the apparent ones which

14. We can postulate the influence of Psychoanalysis, here. The Freudian representation of the psyche, in different levels, where sometimes the conscious level isn't aware of what the individual "really wants" deep down in the unconscious level, seems to be coherent with Schlick's ideas and would help to justify them. The strong influence of Schopenhauer—whose philosophy is extremely similar to Freud's— may support such hypothesis. As Leinfellner (1985, p. 328) notes, "Schlick's concept of the world… is the world of our imagined inner representations plus the sensations of the outer world. It is, therefore, the external world mirrored in our consciousness…. Schlick's concept of the world is heavily influenced by Schopenhauer's philosophy of the world as will and representation. (Schopenhauer's will is an exact synonym of acting and deciding.)".

the agent, or valuer, believes himself to follow. And in so doing he would, perhaps, find it necessary to delve deep into the human soul (Schlick, 1939, p. 20-21).

In short, Schlick's "maneuver" to derive normative statements from descriptive statements can be presented in three steps: (1) it is the case that there is a verifiable fact in nature called "the good"—otherwise it would be a *meaningless* word; (2) even though we cannot access "the good" directly, we can, at least, establish the particular situations in which human beings use this word; (3) since human being use such word in a normative way, scientists, by stating the situations to which "the good" applies, are stating a normative fact.

We can say that Schlick tries to justify his ethics by claiming that the norms of such an ethics are simply prescribing the behaviors that we would select if we really understood the facts about our decision making and our own psychology. Despite its superior structure, Schlick's ethics share the same limitation of Comte's: even if Schlick's psychological project were put in practice and we really could state what someone *would* do in certain "cognitively superior" conditions, we would not be justified to infer that we should *impose* whatever this *hypothetical* cognitively superior person decides over the will of the *actual* person[15]. That the decisions of this cognitively superior person should be considered as morally good is a reasoning completely independent from the scientific research and does not follow from such research.

15. The philosopher Richard Brandt proposed a similar argument. He argues that "the word 'good' should be defined as meaning that a thing is rational to desire, in the sense that one would desire it after a process of 'value-free reflection' that Brandt calls 'cognitive psychotherapy'" (Velleman, 1988, p. 353). The difference is that, in Brandt's cognitive psychotherapy, even though scientific knowledge is presupposed, the "cognitively superior" state can be achieved by the individual's proper reasoning, instead of depending on science alone.

Schlick's project could state "the norm x is normatively imposed over the members of the group y", but never state "the norm x should be imposed over the members of the group y". The *imposition* of these behaviors as norms doesn't follow from such a psychological project, even if it were successful in its empirical dimension.

### 1.3. Harris and the Scientific Ethics in the 21st Century

After addressing a case from the 19th century and a case from the 20th century, one may ask whether or not the dream of a scientific ethics is alive in the 21st century. The influential neuroscientist and philosopher Sam Harris can be addressed here as an example to show that the dream of a scientific ethics is still present in the 21st Century. Harris will be used as an example neither for having provided significant contributions for moral philosophy, nor for the relevance of his work in moral psychology. He will be addressed because he, in a very genuine and passionate way, gives voice to a discourse from moral psychology which is very appealing to scientists and enthusiasts of scientific knowledge nowadays: the discourse that claims that science can tells us what is morally right and what is not. In this section, I'll focus on the arguments instead of the explanations of the background, since the cultural influences leading Harris to search the sources of morality in science are analogous to the ones acting over Schlick and Comte, i.e., the increasing trust in the capacities of science and the decreasing social legitimacy of theology and philosophy.

A relentless critic of religion, Harris believes that "[t]he fact that our ethical intuitions have their roots in biology reveals that our efforts to ground ethics in religious

conceptions of ‘moral duty’ are misguided” (Harris, 2004, p. 172). Similarly to Comte and Schlick, Harris, as a scientist, doesn’t accept the legitimacy of underpinning morality in Metaphysical or Theological theories. The scientific-based discourse is, for Harris, what give us the reliable set of descriptive statements about reality. Like the other authors addressed here, he is compelled to find a way of deriving morality from science. Harris, however, has a tool that wasn’t an option for our previous authors, namely, neuroscience, that he identifies as our best chance to a universal morality.

We can point to similarities between Harris and the previous authors addressed here. In a passage that resembles Comte, Harris says that “[t]here will probably come a time when we achieve a detailed understanding of human happiness, and of ethical judgments themselves, at the level of the brain” (Harris, 2004, p. 175). And, in a passage that resembles Schlick, he says:

> Given that there are likely to be truths to be known about how members of our species can be made as happy as possible, there are almost certainly truths to be known about ethics. To say that we will never agree on every question of ethics is the same as saying that we will never agree on every question of physics (Harris, 2004, p. 182).

In his book *The Moral Landscape* (2010), Harris presents a clear defense of his idea of underpinning morality in science. He is very direct when it comes to explaining his intentions in this book: “I am not [just] suggesting that science can give us an evolutionary or neurobiological account of what people do in the name of ‘morality.’ Nor am I merely saying that science can help us get what we want out of life” (Harris, 2010, p. 28). What Harris wants to make clear, here, is that his work is not just a scientific explanation of our moral inclinations as many studies are (e.g. Renwick Monroe, Martin, & Ghosh, 2009). These studies can even say that “[a]t its core, the moral sense itself will be the same

regardless of cultural variation" (Renwick Monroe et al., 2009, p. 627). But Harris wants more: he claims that science, alone, will discover *moral truths*, independent of our personal values, and will provide us with a scientific morality. He says:

> I am arguing that science can, in principle, help us understand what we should do and should want—and, therefore, *what other people should do and should want* [emphasis added] in order to live the best lives possible. My claim is that there are right and wrong answers to moral questions, just as there are right and wrong answers to questions of physics (Harris, 2010, p. 28).

His argument can be presented, in short, like this: (P1) the exact aim of moral values is the well-being of conscious creatures; (P2) the well-being of conscious creatures is something that can be known scientifically; therefore, (C) moral values can be evaluated scientifically based on their capacity to achieve their goals. Harris also adds that "[t]he most important of these facts are bound to transcend culture—just as facts about physical and mental health do" (Harris, 2010, p. 2). The objective status of such a scientific ethics would justify his optimistic assertion that "[t]here is every reason to believe that sustained inquiry in the moral sphere will force convergence of our various belief systems in the way that it has in every other science" (Harris, 2004, p. 175).

Firstly, we can accept that it seems to be the case that science can inform us how to achieve the goals we want more efficiently: in a banal sense, we can argue that morality can have a 'scientific dimension' in the sense that science can explain how we can achieve more efficiently whatever such morality prescribes. In fact, the conclusion of Harris's argument follows from the premises. But the existence of such a 'scientific dimension' is true for any moral principle, not only for the one presented by Harris. In the same way in which Harris says that "[i]f there are objective truths to be known about human well-being… then science should one day be able to make very precise claims about which of

our behaviors and uses of attention are morally good" (Harris, 2010, p. 8), we could say, after establishing that 'friendship' is all that matters morally, that "if there are objective truths to be known about human *friendship*, then science should one day be able to make very precise claims about which behaviors and uses of attention are morally good". Both conditionals, here, are obviously true statements. But it would be a mistake to infer from these statements that our identification between friendship and morality or our identification between well-being and morality is itself based in science.

While the argument used here to sum up Harris's reasoning is valid, the main problem is the assumption of (P1) ("the exact aim of moral values is the well-being of conscious creatures"). Harris identifies the goal of moral values with well-being, and, from this, he derives his ideas about how science can help us to achieve this goal. Such identification is problematic in three ways.

The first problem with (P1) is that the acceptance of the moral principle—in this case "well-being"—is *not* derived from descriptions. Harris like to use the case of corporal punishment in some American schools (see Harris, 2010, p. 3-4) as an example of what we know "scientifically" to be wrong. He asks: "[i]s there any doubt that this question [about corporal punishment] *has* an answer?" (Harris, 2010, p. 3). And, mentioning research that shows that corporal punishment is disastrous in education, he concludes that this is an issue we can scientifically answer. But the problem is that, without the assumption that children must learn or that we should take care of them, we would never know scientifically what is right. We would simply know that "children learn better if we don't beat them"—which is simply a description as moral as any other description of a natural fact. We can derive

norms from previous norms scientifically[16], but this happens only if we already have accepted normative principles before. What Harris wants, however, is the "starting point": he claims that we can find "moral truths" through science. He says: "[i]f moral truths transcend the contingencies of culture, human beings should eventually converge in their moral judgments" (Harris, 2010, p. 179). The ambition of *finding* pure moral truths through science is defended by Harris using arguments that only indicate the *derivation* of a norm from another.

The second problem with (P1) is that *even if we could* know that our brains are hard-wired to promote "well-being of conscious creatures"—whatever that would mean—, the question of whether our morality should adopt such a tendency as a norm would remain, again, an open question. The fact that our brain structure is such that it makes us act in given ways or pursue given goals should not be accepted as a sufficient reason for us to adopt or prescribe such behaviors—this would clearly be a naturalistic fallacy.

The third problem with (P1) is that "well-being" is a vague term. Harris himself admits not being sure about what it is (see Harris, 2010, p. 8) and accepts that it changes through time. Even more problematic than the concept of "well-being" is "well-being of conscious creatures", since given actions would be ambiguously related to it, i.e., it would benefit some conscious creatures and harm others. To adopt "well-being of conscious creatures" as our moral principle would lead to philosophical difficulties, like the well-being of the actual individuals vs. the well-being of the future individuals, the well-being of humans vs. well-being of other conscious creatures, the well-being of "the experiencing

16. E.g., if we accept the norm "you should not kill", and science informs us that the substance x is poison, we can infer, based in science, that "we should not put the substance x on someone else's food".

self" vs. the well-being of the "remembering self" (see Harris, 2010, p. 184), etc. The old question "what is moral?" would be replaced by "what is the well-being?", keeping the same plurality of discussions about the proper definition[17].

The vagueness of "well-being" also makes impossible the *falsification* of Harris's hypothesis of a scientific morality. He says: "[c]learly, if there is a more important source of value that has nothing to do with the well-being of conscious creatures (in this life or a life to come), my thesis would be disproved" (Harris, 2010, p. 189). But, since well-being is a vague term, there is no way to falsify his hypothesis.

It seems that a frequent problem for the proponents of a scientific ethics is the naïve representation of the *modus operandi* of what Schlick called a "moralist", i.e., someone who prescribes moral rules. The risks involved in the idea of establishing norms for human behavior are perhaps obliterated both by their trust in science and by their superficial account of how this moralist thinks. Schlick considered that the philosopher who "is in a real opposition to those final norms recognized by life… has unwittingly become a moralist, that… feels uncomfortable in the role of a knower and would prefer to be a creator of moral values" (Schlick, 1939, p. 20). If a moralist is deliberately arbitrary, then the scientist would never be a moralist. However, a more realistic account of moralists would recognize that moralists (e.g., Plato, Augustine, Aquinas, etc.) are doing something

17. A less charitable reading of Harris could claim that, ultimately, Harris's argument for using "well-being" is a case of *ad ignorantiam* fallacy. Says Harris: "[m]uch of the skepticism I encounter when speaking about these issues comes from people who think 'happiness' is a superficial state of mind and that there are far more important things in life than 'being happy.' Some readers may think that concepts like 'well-being' and 'flourishing' are similarly effete. However, I don't know of any better terms with which to signify the most positive states of being to which we can aspire" (Harris, 2010, p. 183).

analogous to what Schlick or Harris are doing, i.e., recognizing, through reason, some facts (descriptions) and arguing that some rules (prescriptions) follow from these facts.

Harris's scientific morality tries to justify its normativity by saying that it is just a condition to achieve what people *really* want. Instead of the salvation of the soul or the higher pleasure, he uses the notion of "well-being", and, as it is in other moral doctrines, it is precisely in the definition of the "object of morality" (Harris's well-being, here) that the scientist become a moralist, being subjected to personal inclinations.

While Schlick argued against the is-ought problem by providing a different notion of normativity, Harris argues against the division between facts and values. He states that "the divide between facts and values is illusory in at least three senses" (2010, p. 11). I'll explain these three points below.

The first point is that "whatever can be known about maximizing the well-being of conscious creatures… must at some point translate into facts about brains and their interaction with the world" (2010, p. 11). What Harris does here is the same as saying "Poseidon's wrath is a physical phenomenon", after defining "Poseidon's wrath" as a "storm". I.e., he defines "values" as something factual and then says that it is factual. But, as said before, the 'moral facts' that Harris claims that science can show us (e.g., that corporal punishment is wrong) are dependent on previous moral values. In other words, we still need a moral 'starting point', based in what we would derive other norms scientifically. As explained in the previous sections, it is impossible to derive any normative statement from science's descriptive statements. Therefore, Harris's example doesn't explain the origin of moral values, at most it can show that some prescriptions are derived from others,

and that such derivation can be done scientifically. But he leaves untouched the question about the source of our initial moral values.

Harris's second point against the division between facts and values is that "the very idea of 'objective' knowledge… has values built into it, as every effort we make to discuss facts depends upon principles that we must first value (e.g., logical consistency… parsimony, etc.)" (2010, p. 11). An easy answer to this is that the influence of cognitive values is different from the influence of non-cognitive values, as explained before. But this is a complex topic which goes beyond the scope of this thesis. But we can say that the fact that science is influenced by *some* values doesn't mean that it can be influenced by *any* value.

Finally, in the third point, Harris argues that "beliefs about facts and beliefs about values seem to arise from similar processes at the level of the brain: it appears that we have a common system for judging truth and falsity in both domains" (2010, p. 11). And he adds:

> This finding of content-independence challenges the fact/value distinction very directly: for if, from the point of view of the brain, believing "the sun is a star" is importantly similar to believing "cruelty is wrong," how can we say that scientific and ethical judgments have nothing in common? (Harris, 2010, p. 122).

Even if Harris's research on beliefs (see Harris, 2010, p. 120) is supported by facts and we share the same cognitive processes when evaluating facts and values, we cannot infer that science and philosophy should share our brain limitations. In fact, science is precisely a way to avoid the mistakes from our flawed and biased natural reasoning. The area of the brain in which such domains are processed isn't saying anything about the limitations of the concepts themselves.

We can find in Harris the same "maneuver" presented in the previous sections. The scientist affirms that the rules x and y are objectively the ethically right ones to adhere to, and, if someone thinks otherwise, it is so because this person is ignorant of them—the rules x and y are *what we would want if* we had a proper access to scientific knowledge. Thus, in the same way in which we correct someone who states that time is absolute, expecting that this person would change this false belief, we would correct someone acting differently by using the standards of what the scientists state to be scientifically good.

The appeal of a scientific ethics is clear. Instead of a spiritual power legitimizing our morals, we would have the clean, objective, enlightened power of science legitimizing it. The normativity that obligates us to believe in the Theory of Relativity would obligate us to act in given ways. As Harris says, "[s]aving a drowning child is no more a moral duty than understanding a syllogism is a logical one" (Harris, 2004, p. 172). The same justification we have for dismissing a mistaken statement about nature would be possible regarding ethics: "[w]hen was the last time that someone was criticized for not 'respecting' another person's unfounded beliefs about physics or history? The same rules should apply to ethical, spiritual, and religious beliefs as well" (Harris, 2004, p. 176). But, unfortunately, the arguments for such ethics are not enough to support his claims, as I presented here.

A possible defense of Harris's moral project may be to accept that its premises are a *choice*. One can decide to adhere to a morality that is defined by what the main scientists happen to believe. Indeed, at this point in the 21st century, the condition of the coherence between our morality and our scientific facts may be a fair requirement.But the believer would still have some problems: it would require a consensus of scientists about morality or an arbitrary election of whose opinions would count as the "opinions of science". If

someone wants to derive “moral authority” from “scientific authority”, we shouldn’t consider it to be significantly different from a metaphysical or a religious position. But what needs to be clear is that such an ethical project would *not* be a science-*based* ethics (as something that follows from scientific knowledge). It would be just a science-*informed* ethics—and it would not be significantly different from other ethics, sharing their limitations and their arbitrariness. In short: the choice to pick a scientifically described phenomenon to count as the moral good doesn’t make this choice more scientific than the decision to pick a theologically described rule to count as the moral good would makes this decision more divine.

# Chapter 2

# The Science of an Ethical Dream

If looking at the history of the altruism and morality debate in biology teaches us anything it is that our attempts to understand these phenomena in nature have been particularly susceptible to insinuations of how we wish to see ourselves in the world (Harman, 2014, p. 161-162).

In the last chapter, I addressed some philosophical arguments for the idea of deriving ethics from science. In this chapter, I will focus on the scientific study of what can be an innate unselfish concern with others. The question underlying the following chapter is "how moral values can influence scientific research?" Such inquiry will have two sections: one will be a historical case and the other an investigation of scientific terminology in contemporary science.

Firstly, I'll present an historical example from evolutionary biology in order to explore the question about how non-cognitive values can influence scientific research. This section will address the work of Peter Kropotkin, an evolutionary biologist from the 20$^{th}$ century, who argued against the notion of nature as a "struggle for existence". Instead, he claimed that nature is mostly cooperative. The case is particularly interesting due to the comparison between Kropotkin and another evolutionary scientist, T. H. Huxley, whose views of nature were precisely the opposite of Kropotkin's. How could two respected

scientists, sharing a Darwinian background, reach such radically opposed conclusions? This will be one of the main questions in this section.

The second case study will be about the term “altruism”. From all the concepts used to link ethics and science, altruism certainly is the most common. Since Comte, the idea of an altruistic nature feeds the imagination of those who want to argue for morality in nature. This second section will explain the different usages and limitations of the term in biology and psychology.

### 2.1. Kropotkin and a Cooperative Natural Selection

Undoubtedly, one of the reasons for the astonishing popularity of Charles Darwin’s theory of evolution is how easily it allows anyone, after a brief introduction to the theory, to imagine possible explanations for almost every behavior of living organisms. Once one looks at the complexity of nature through the ‘glasses of the evolution’, the chaotic plurality of nature slowly turns into a coherent whole that follows from simple rules. However, some behaviors seem to resist to fit in one’s *prima facie* picture of nature, compelling the observer to a more complex explanation. One of the puzzling behaviors that sound dissonant to Darwin’s theory of evolution is altruism.

Darwin himself, while investigating the altruistic behavior of neuters or sterile females in insect-communities, acknowledged the difficulty concerning the apparent incompatibility of altruism and natural selection: “[altruism] at first appeared to me insuperable, and actually fatal to my whole theory” (Darwin, 1859, p. 214). As Sober & Wilson (1998, p. 3) explain this apparent difficulty: “[t]he basic idea of natural selection is

that characteristics evolve because they help the individuals who possess them to survive and reproduce", therefore, it is intuitive to infer that "helping other individuals to survive and reproduce at the expense of one's own survival and reproduction is the very thing that natural selection will eliminate". Altruistic behaviors are a clear disadvantage for the performer, thus, they were supposed to be eliminated by natural selection. The idea of a selfish or self-regarding organism appears more at home in Darwin's theory. As Wilson (2015, p. 4) states,

> Altruism occupies center stage in Darwinian thought because it appears difficult to explain as a product of natural selection. If natural selection favors traits that cause individuals to survive and reproduce better than other individuals, and if altruistic acts increase the survival and reproduction of others at a cost to the altruist, then how can altruistic traits evolve?

The answer to the question above was given by Darwin in terms of what is known as "kin selection". For Darwin, the problem of altruism in nature, "though appearing insuperable, is lessened, or… disappears, when it is remembered that selection may be applied to the family, as well as to the individual" (Darwin, 1859, p. 215). As Harman (2014, p.148) explains, "Darwin solved the conundrum to his satisfaction by positing the notion of a 'community': Who benefits from the toil of the nursemaid, the forager, and the soldier? The queen… and by extension the entire growing family". In other words, for Darwin, the mystery of altruism in nature can be solved once the *unit of selection*—i.e., the unit over which natural selection's rules will be applied—is the family rather than the individual[18].

18. Some contemporary authors consider that colonies should be considered as a super-individual, or, as Haber (2013) argues, simply an individual like any other.

Nevertheless, some authors of evolutionary biology's tradition thought that the explanation in terms of benefiting one's relatives wasn't enough. In the beginning of the 20th century, the Russian scientist and philosopher Peter Kropotkin defended the idea that blood kinship, alone, couldn't explain every case of cooperation in nature. Exploring the Eastern Siberia and the Northern Manchuria as a scientist, Kropotkin witnessed an environment where cooperation and mutual support were the rule. He published the results of these journeys in his *Mutual Aid: A Factor of Evolution* (1902), a book inspired, in part, to dispute an article written by T. H. Huxley, "The Struggle for Existence in Human Society" (Huxley, 1902, p. 195-236), where Huxley pictures a nature radically distinct from that described by Kropotkin. A brief explanation of Huxley's thought will be useful to understand Kropotkin.

Huxley believed that nature was intrinsically cruel and competitive. He states that "[i]f we confine our attention to that aspect which engages the attention of the intellect, nature appears a beautiful and harmonious whole", but if we look to nature from a more realistic point of view, "then our verdict… can hardly be so favourable" (Huxley, 1902, p. 195). For Huxley, altruism without kinship, just like every ethical behavior, was a fiction created by humans. For him, as Harman (2014, p. 150) mentions, "[t]he human brain, after all, a product of evolution as much as the feathers of the peacock… could transcend the natural imperative, replacing indifference and necessity with caring and ethical progress". Through moral education, humans could develop the altruistic behavior in the absence of kinship relations, but it was nothing more than an *exception* in nature.

The metaphor of "struggle for existence", used in the title of Huxley's essay is emblematic of the Huxley's and Kropotkin's radically distinct approaches—which, in turn,

is emblematic of the divergence between the evolutionists from the British and the Russian traditions. The metaphor was firstly used by Darwin himself, who stated that "[n]othing is easier than to admit in words the truth of the universal struggle for life" (Darwin, 1859, p. 63). The inspiration for such metaphor comes from the work of Thomas Robert Malthus.

> The Reverend Thomas Robert Malthus… published *An Essay on the Principle of Population, as It Affects the Future Improvement of Society*, a short, trenchant critique of Godwin, Condorcet, and other advocates of the Enlightenment view that humans and human society were infinitely perfectible. For Malthus such optimism foundered on an inexorable natural law: "Population, when unchecked, increases in a geometrical ratio. Subsistence increases only in an arithmetical ratio." Thus, all organisms were subject to "a strong and constantly operating check on population." Among plants and animals this check took the form of "waste of seed, sickness, and premature death." Among humans it was expressed in "misery and vice." (Todes, 1989, p. 13).

In an explicit reference to Malthus, Darwin (1859, p. 64) stated:

> Hence, as more individuals are produced than can possibly survive, there must in every case be a struggle for existence, either one individual with another of the same species, or with the individuals of distinct species, or with the physical conditions of life. It is the doctrine of Malthus applied with manifold force to the whole animal and vegetable kingdoms.

But while "[f]or Darwin and other leading British evolutionists, the expression 'struggle for existence' appealed to common sense, and its Malthusian associations posed no problem", the Russian evolutionists "reacted negatively to what they perceived as a transparent introduction of Malthusianism—or, for some, simply the British enthusiasm for competition—into evolutionary theory" (Todes, 1989, p. 3). The idea of struggle for existence, for the Russian evolutionists, was, ultimately, a way of threating human competition as the law of nature. For Kropotkin and the other members of the Russian school of evolutionary biology, the *values* underlying the usage of such Malthusian metaphor were not cognitive values. The metaphor was not used due to its capacity to

simplify an observable pattern of nature—what would be justified as something motivated by cognitive values—, instead, it was motivated by the values adopted by the British authors, influenced by the Victorian culture and ideas such as Eugenics.

With the conception of nature as struggle for existence, Huxley considered that morality was a thin veneer covering the "real nature" of humans, which is a violent and egocentric Hobbesian beast, always threatening from underneath. The Huxleyan nature was a cruel and amoral force in which the "Hobbesian war of each against all was the normal state of existence" (Huxley, 1902, p. 204). The following passage makes explicit how radically violent Huxley's picture of nature was: "[f]rom the point of view of the moralist the animal world is on about the same level as the gladiator's show. The creatures are fairly well treated, and set to fight—whereby the strongest, the swiftest, and the cunningest live to fight another day" (Huxley, 1902, p. 199-200).

In his book *Mutual Aid: A Factor of Evolution*, Kropotkin attempts to counter Huxley's ferocious conception of nature through a myriad of examples, from animal behavior to historical events, that he considered to be evidence of nature's "tendency" to cooperation. Opposing Huxley, Kropotkin stated:

> [W]herever I saw animal life in abundance… I saw mutual aid and mutual support carried on to an extent which made me suspect in it a feature of the greatest importance for the maintenance of life, the preservation of each species, and its further evolution (Kropotkin, 1902, p. xii).

Naturally, Kropotkin didn't deny the existence of what he called "an immense amount of warfare and extermination going on amidst various species" (Kropotkin, 1902, p. 5). Competition was a fact and he even explicitly rejected romanticized views of nature such as Rousseau's (Kropotkin, 1902, p. 4). However, despite the existence of struggle in

nature, Kropotkin believed, against Huxley[19], that "mutual aid is as much a law of animal life as mutual struggle, but that, as a factor of evolution, it most probably has a far greater importance" (Kropotkin, 1902, p. 5).

Inspired by the zoologist Karl Kessler, who argued that mutual aid was a mechanism more efficient for natural selection than competition (see Kropotkin, 1902, p. 7), Kropotkin thought that cooperation and altruism were not simply ways of protecting one's offspring, but the standard behavior selected by natural selection. As Dugatkin (2007, p. 1373) explains, Kropotkin "believed that he saw altruism in every species that he came across, and all this altruism, as Kropotkin saw it, was divorced from blood kinship". Instead of a bizarre exception in natural selection, cooperation and altruism were, for Kessler and Kropotkin, behaviors that increase the fitness of individuals. In Kropotkin's words:

> The animal species, in which individual struggle has been reduced to its narrowest limits, and the practice of mutual aid has attained the greatest development, are invariably the most numerous, the most prosperous, and the most open to further progress" (1902, p. 242).

In a famous passage, Kropotkin says: "'Don't compete! — competition is always injurious to the species, and you have plenty of resources to avoid it!' That is the *tendency* of nature" (1902, p. 61). He argued, thus, that cooperation is not motivated by or dependent on some sort of morality: it is a natural instinct. And it is not so because nature is good or moral but because mutual aid is more *efficient* from an evolutionary perspective, or, in his words, cooperation is "the surest means for giving to each and to all the greatest safety, the

19. Perhaps, one could be inclined to add "and against Darwin, too!", but Kropotkin believed that Huxley's interpretation of Darwin's work was biased. Kropotkin believed that Darwin actually had some intuitions pretty close to what he defends, but never fully developed such ideas. Huxley's description of a Hobbesian nature was, for Kropotkin, just the belief of the school of writers, lead by Huxley himself, who took "possession of Darwin's terminology rather than his leading ideas" (Kropotkin, 1902, p. 63).

best guarantee of existence and progress, bodily, intellectual, and moral" (Kropotkin, 1902, p. 61).

Even though cooperative behaviors were not caused by moral preferences, Kropotkin argues that, since nature is cooperative, some moral principles *follow* from nature. Linking nature and morality, he states:

> That mutual aid is the real foundation of our ethical conceptions seems evident enough. But whatever the opinions as to the first origin of the mutual-aid feeling or instinct may be—whether a biological or a supernatural cause is ascribed to it—we must trace its existence as far back as to the lowest stages of the animal world. (Kropotkin, 1902, p. 246).

For him, morality could follow naturally from our instincts, making the intellectual morality from the Church or from philosophy—what would be *artificial* moralities—something superfluous for an ethical society. The following passage shows Kropotkin's idea of a natural morality in a very explicit way.

> Ask the ants if it would be right to refuse food to other ants of the same ant-hill when one has had one's share.… Or again, ask the sparrows living in your garden if it is right not to give notice to all the little society when some crumbs are thrown out, so that all may come and share in the meal.… Finally, ask primitive man if it is right to take food in the tent of a member of the tribe during his absence.… The ant; the bird, the marmot, the savage have read neither Kant nor the Fathers of the Church… yet all have the same idea of good and evil. And… what is considered as good among ants, marmots, and Christian or atheist moralists is that which is useful for the preservation of the race; and that which is considered evil is that which is hurtful for race preservation. The idea of good and evil has thus nothing to do with religion or a mystic conscience. It is a natural need of animal races. And when founders of religions, philosophers, and moralists tell us of divine or metaphysical entities, they are only recasting what each ant, each sparrow practises in its little society (Kropotkin, 1970, p. 90-91).

What is particularly interesting about the passage above is that it wasn't taken from Kropotkin's scientific work. It was part of a pamphlet he wrote, called "Anarchist Morality" (1970, p. 79-113), where he defended his political views. At this point, considering the examples of the first chapter, it is tempting to explain Kropotkin's

conception of a cooperative nature as a consequence of his personal inclinations, which appreciate mutual aid and a society freed from government and moral codes. How can we be sure that Kropotkin's scientific statements about nature aren't determined by his political views? In other words, perhaps Kropotkin's scientific idea of an altruistic nature was not only correlated, but *caused* by his political views—i.e., by his non-cognitive values. I'll call this hypothesis "H1". In fact, H1 would not only conveniently fit in this thesis, cohering with the previous chapter, but would also help to explain why his perspective differs so dramatically from Huxley's.

Kropotkin's answer to this problem is simple. He would agree that it is extremely difficult to distinguish the scientist Kropotkin from the anarchist Kropotkin. But he would add that such difficulty is not a problem. For him, no demarcation was really necessary in this case, since he believed that his Anarchism was supported by the scientific facts. Kropotkin observed in nature what he believed to be the basis of Anarchism. Consequently, he could claim that his Anarchism was science-based—not a simple ideology, but the very rule of nature. If nature is essentially altruistic, there would be no need for strong hierarchical and moral rules: our "sense of equality" (Kropotkin, 1970, p. 99) could be our ruler. As Harman (2014, p. 150) comments, Kropotkin believed that it was "the return to animal origins that promised to save morality for mankind". Thus, he would say that, if there are a relation of *causation* between his political and scientific views, it would be the political that is caused by the scientific—not the other way around.

Kropotkin's answer in the paragraph above wouldn't really solve the problem. Our suspicion, here, is not about the philosophical reasoning and how it *uses* scientific facts (as in Harris section, for example), but rather a suspicion about the *scientific facts themselves*.

After reading Kropotkin's explanations, we could still wonder whether the very observations he described as a scientist were influenced by his political views. We could still ask: how can we be sure that Kropotkin's observations of nature were not *biased* and showed him just what he wanted to see?

In Kropotkin's defense, we can point out that, as Todes (1989, p. 123) says, "Kropotkin first questioned Darwin's approach to the struggle for existence while exploring Siberia as a youth and was an accomplished and celebrated naturalist years before his political views crystallized". In addition to this, Kropotkin's ideas were not controversial: "his ideas about cooperation in nature were quite common among Russian naturalists of varying political perspectives" (Todes, 1989, p. 123). Thus, considering these two facts mentioned by Todes, the idea that Kropotkin's ideas are simply a way to support his political views doesn't seems to be a strongly supported claim, since (1) his ideas could have been independently acquired and (2) they were not his invention, but a development of what was a common-sense of his school[20]. Secondly, what could corroborate the hypothesis that Kropotkin's scientific work was, in fact, a proper scientific research is the fact that some of his intuitions about an innate inclination to morality are still debated nowadays, and have actually found some support in recent studies (e.g. Renwick et al., 2009; and Barragan, & Dweck, 2014). Therefore, as Todes states, even if "Kropotkin's theory of mutual aid certainly had an ideological dimension", it doesn't mean that his work

20. As said before, Huxley's perspective was not free from cultural influence as well: "for Kropotkin, Huxley's article reflected a broad consensus among English intellectuals" (Todes, 1989, p. 132). Scientifically, both Huxley's and Kropotkin's views reflected mainstream attitudes and theories of their respective intellectual environments.

can simply be “dismissed as the idiosyncratic product of an anarchist dabbling in biology” (Todes, 1989, p. 123).

We can, therefore, conclude that the *convergence* between Kropotkin’s political and scientific views cannot be answered with the easy accusation of an influence of his political views over his scientific work as H1 states. In fact, such accusation would be, after all, *ad hominem* (see Sober & Wilson, 1998, p. 17). However, regardless of the explanations above and the abandonment of H1, a question still unanswered: how could two respected evolutionary scientists (Kropotkin and Huxley) reach such radically opposite conclusions based in the same theory?

Lee Dugatkin (2011, p. 12-36) offers an intriguing explanation of this conflict. He argues that these differences can be accounted for through the context and locales of their investigations. While Huxley, just like Darwin before him, observed nature primarily in the tropics, Kropotkin and the Russian biologists observed the completely different environment of Siberia and Northern Manchuria. The key difference that could explain the disparities between their views is that Kropotkin’s nature was the struggle against *underpopulation*, and, in this context, cooperation tends to be advantageous, while Huxley’s nature was the struggle against the *overpopulation,* where direct competition is more occurrent. These differences were decisive for their divergence.

Clearly, while Huxley’s environment was easily describable in Malthusian terms—this would not be the case for the cold underpopulated Russia. Such difference may explain why nature as “struggle for existence” was a common perspective for the English intellectuals and non-sense for the Russian intellectuals. Dugatkin (2011) also mentions a secondary factor that may have accentuated the differences between Kropotkin’s and

Huxley's perspectives: while Kropotkin had been raised in a wealthy context, Huxley faced many more economic challenges.

Even though we cannot affirm H1, this presentation of the Kropotkin-Huxley debate shows how scientific research can be subjected to contingent factors. In other words, even if we cannot affirm that Kropotkin's scientific research was influenced by his political views, we cannot affirm that it wasn't. The possibility of non-cognitive values actively influencing their views is enough for our skepticism over the scientific status of their work.

To conclude this section, I'll offer another hypothesis to explain this debate. Instead of our naïve H1, we can consider another hypothesis: what allowed the radical divergence between Kropotkin and Huxley was the interpretation of *contingent* behaviors (cooperation, in the case of Kropotkin, and struggle for existence, in the case of Huxley) as if they were *necessary* behaviors. In other words, instead of considering a set of behaviors as contingently beneficial, they considered something always beneficial for every organism. Based on this mistake, Kropotkin inferred that altruistic behaviors were fixed by evolution at the core of our instinctual nature, while Huxley inferred that selfish behaviors were our real inclination. Both authors are making wrong inferences even if their scientific observations were correct.

### 2.2 The Concept(s) of Altruism

One of the problems underlying the Kropotkin-Huxley debate is the language used by both authors. Perhaps as much as the contextual differences between them, the

imprecision of the terms used can be taken as a major factor of their disagreement. “Cooperation”, “mutual aid”, “struggle”, etc., are all metaphorical and/or vague terms.

The study of animal behavior, especially when the object of study is the behavior under natural conditions, is easily vulnerable to the scientist’s bias. We can identify this vulnerability as a consequence of the lack of control over the variables of the natural environment and the limitations of the observation itself[21]. Many behaviors can be misinterpreted or skip the scientist’s perception. Even when a tool such as an ethogram[22] is available, there will be always a limited number of behaviors catalogued, which restricts the scientist’s perception as well (see Mech & Boitani, 2010, p. 40-42).

Considering these limitations inherent to ethology, the usage of terms that don’t represent a specific measurable property makes virtually impossible the avoidance of the influence of non-cognitive values in scientific research. In this section I’ll explore some ways in which contemporary science tried to solve some of the issues raised on the debate about cooperation in nature and point out some of the problems involved in this process. My study case will be the concept of altruism, which is a key term in this debate.

The first point that needs to be clear is that the claim that there is a pattern in nature called “altruism” (as a natural kind), and that philosophers and scientists discuss how to

21. These two problems are particularly evident when we want to make inferences about non-human animals’ mental states, e.g., meta-cognition. Cognitive ethologists are far from reaching the same scientific consensus compared to branches of biology and psychology that don’t share these problems. While some authors, as Carruthers (2008, p. 58), criticize the attribution of meta-cognition to animals, holding that “there is no need to postulate meta-cognitive processing in order to explain the data”, on the other extreme, researchers like Cabanac, et al., (2009) want to defend the idea that even single cell organisms can be conscious.

22. A catalogue of behaviors from a given animal, used by ethologists.

better describe it, is false. As we will see, altruism has had different meanings through its existence, going from mental states to the number of offspring. There isn't an unchanged essence or a fixed reference or extension for the term altruism. As Sober & Wilson (1998, p. 17) state, the difference between definitions of altruism in science is "more than simply looking at the same question from a different angle; rather, the phenomenon under study undergoes a transformation as well". That said, we can abandon the idea of a "conceptual analysis" as some kind of inquiry on stable and constant mental/physical object/structure. Here, we will cover different altruisms, explaining the limitations and benefits of some of these definitions.

Before addressing the scientific definitions of altruism, it will be interesting to make explicit the vagueness of the standard definition of altruism as it is used in the common-sense. The list below is not exhaustive, of course. It just shows a few examples of vagueness in the standard notion of altruism. Consider the following definition:

> *Altruism occurs when an action (a) is performed by an individual who has both a desire/intention to benefit another individual through the action (a) and a belief which states that the action (a) will not result in any benefit to the performer.*

The following list shows some imprecisions involved in this definition: (1) it is not clear what would count as a "benefit"; (2) it is not clear what are to be considered, the immediate results of the action or the results in terms of a lifetime; (3) it is not clear whether

the actual benefit to someone else is required or if decreasing the probability of loss[23] would be enough; (4) it is not clear whether the performer's belief that the action (a) will not bring a benefit should be a true belief; (5) it is not clear if avoiding the performer's *suffering*—that would follow if the performer didn't act—would count as a benefit for the performer; (6) it is not clear if the other individual should be someone unrelated to the performer or not[24]; (7) it is not clear which are the limits to what we can consider as another individual[25]; etc.

When it comes to the plurality of scientific definitions of altruism, we can recognize one major division. On one side, we have the *psychological* altruism, and, on the other, the *biological* altruism[26]. In this section I'll address both of them, exploring their benefits and limitations.

Psychological altruism is the most intuitive form of altruism. It is almost identical with the conception of altruism used in the everyday life. As Clavien & Chapuisat (2013, p. 126-127) state, "[i]n everyday language, altruism occurs when individuals are disposed to sacrifice part of their personal interest in favour of others; it is an honourable gift given without any expectation of future personal reward". Using more technical terms, we can

23. E.g., taking someone away from a dangerous place.

24. Since helping the ones that the performer loves could be a selfish indirect action.

25. We can ask whether the other needs to be a human and even a living organism. One could ask with Sober & Wilson (1998, p. 229): "what about people who care about a nation, a religion, an ethnic group, or a cultural tradition, not just as means but as ends in themselves?... are they altruistic?".

26. It is important to point out the fact that there are many psychological altruisms and many biological altruisms. However, for the purposes of this thesis, this broad distinction will be sufficient. For an appreciation of the plurality and complexity of definitions of altruism, see Clavien, & Chapuisat (2013).

say that psychological altruism is the psychological disposition to perform behaviors which have the goal of increasing other's welfare without doing so because of further benefits to one's own welfare. The psychological altruist increases the other's well-being *for the other's sake*, not because of some personal benefits that could follow directly or indirectly from such action. As we presented in the last chapter, Comte coined the term altruism as the opposite of egoism, and psychological altruism preserves this original meaning. I'll address psychological altruism in depth later in this section.

Biological altruism, on the other side, is not about what happens in the performer's mind. The motivation, the beliefs, and the desires of the performer are irrelevant. What matters, here, are the observable consequences of the behaviors performed by the actor. What determines if a behavior is biologically altruistic is the effect of such behavior over the number of offspring, both from the performer and from the recipient. As West, et al. (2007, p. 416) define it, biological altruism is "a behaviour which is costly to the actor and beneficial to the recipient; in this case…, cost and benefit are defined on the basis of the lifetime direct fitness consequences of a behaviour". Thus, while it is not clear what would count as a "benefit" in psychological altruism, in biological altruism we have a clear measurable unit—biological altruism is a quantitative concept.

It was the work of W. D. Hamilton that, for the first time, offered the possibility to evaluate altruism quantitatively by establishing something *measurable* to identify altruistic behaviors, making biological altruism possible. Firstly, in his work *The Genetical Theory of Social Behavior I and II* (1964), Hamilton defended a different understanding of the concept of fitness. As Harman (2014, p. 157) explains:

[C]lassical population geneticists had defined 'fitness' as the measure of an organism's reproductive success—the more offspring an organism sired, the greater its fitness. A

corollary of this definition was that the persistence of any behavior or gene responsible for it which reduced an organism's fitness would be difficult to explain.

Going beyond personal fitness, Hamilton coined the concept of "inclusive fitness" (1964, p. 8). The classical definition of fitness considered the individual as the object of selection, but Hamilton focused on the natural selection of the *gene*. Inclusive fitness was "the effect of one individual's actions on everybody's numbers of offspring … weighted by the relatedness" (West, et al., 2007, p. 416). If a behavior is, ultimately, a mean for the gene to spread itself, the offspring of the performer is not essential, since an action of this performer could benefit the offspring of other individuals with the same gene. "[T]he key was adopting a 'gene-eye' point of view.… From the point of view of the gene trying to make its way into the next generation, it made absolutely no difference in whose body it was being carried" (Harman, 2014, p. 158). Therefore, the apparent incompatibility of altruism with natural selection is a consequence of a mistaken perspective that considers the individual as the object of selection instead of the genes.

Perhaps the main strength of Hamilton's argument is a mathematical formulation he developed, that allowed altruism to be evaluated quantitatively[27]. Such formulation was only possible due to a tradition started before him by the fathers of mathematical population genetics, namely, B. S. Haldane and Ronald Fisher, with the mathematical formulation of

27. The passage from a vague concept to a quantitative concept is what Carnap called "explication" (see Carnap, 1950, p. 1-19). However, as we will see later, the reduction of psychological altruism to biological altruism doesn't seem to be justified.

kin selection[28], and Sewall Wright, with the quantitative concept of "coefficient of relationship"[29].

Using Wright's "coefficient of relationship," r, as his measure of genetic relatedness, Hamilton added in the costs (c) and benefits (b) of altruism to his model. The importance of adding b and c to his model was that it allowed Hamilton to take an economic approach to how natural selection might maximize fitness and still allow for the evolution of altruism (Dugatkin, 2007, p. 1378).

Hamilton's formulation, known as Hamilton's Rule, is represented in the formula "$r. B > C$". As Harman (2014, p. 158) explains:

Every altruistic act …would entail both a fitness cost to the altruist …, and a fitness benefit to the receiver …. What Hamilton showed was that, for every social situation, if the benefit (B), devalued by the relatedness between the two actors (r), was greater than the cost (C), genes that play a role in bringing about altruistic behavior could evolve. The greater the relatedness, the greater the chance for, and scope of, altruism.

Biological altruism is a safe definition for altruism: what we mean by benefit is something measurable. The advantages that a quantitative concept of altruism brings for the scientific research are clear. However, the very same property that makes biological altruism advantageous end up being its limitation: this concept is too narrow to evaluate many behaviors that we usually consider to be cases of altruism.

Patricia Churchland (2011) express her discontentment with the term "cooperation" in its standard biological definition[30]. She says that:

---

28. Kin selection is the "process by which traits are favoured because of their beneficial effects on the fitness of relatives" (West, et al., 2007, p. 416).

29. A measure of the degree of consanguinity.

30. Notice that cooperation, here, is not a vague term. It means precisely "a behaviour which provides a benefit to another individual (recipient), and which is selected for because of its beneficial effect on the recipient" (West, et al., 2007, p. 416). The benefit, here, is the very same mentioned for biological altruism, i.e., the lifetime direct fitness

> When my neighbor and I engage in a joint tractor-repairing effort because it is advantageous for both of us and extremely difficult to achieve singly, this would commonly be called cooperation. Nevertheless, because joint tractor-repairing behavior presumably is not the outcome of natural selection (our brains did not evolve to repair tractors), then by the biologist's definition, our venture does not qualify as cooperation (2011, p. 67).

Just like the biological cooperation, biological altruism is not efficient in some cases. Moreover, our common-sense notions of cooperation and altruism seem to be efficient terms to be used in such cases. Considering this, the *reduction* of psychological altruism to biological altruism isn't justified. Assuming biological altruism as the *explication* of psychological altruism, in the Carnapian sense, would be a mistake. As Sober & Wilson (1998, p. 330) sum up, "evolutionary altruism and psychological altruism must be evaluated separately"—they are different concepts, created to solve different problems.

Even if we recognize the superiority of biological altruism as a scientific concept, we need to accept that some cases seem to be better addressed by psychological altruism. But from this fact we cannot infer that psychological altruism is a good scientific concept. I'll address, now, some of the limitations of the psychological altruism when it is used in science. Independently from the definitional vagueness[31], there are *intrinsic* problems with psychological altruism. In short, these problems are (1) the assumption of the mental states described by what is called "Folk Psychology"; (2) the irrelevance of observable behaviors

---

consequences of a behavior. Therefore, the limitation of the concept of cooperation, presented by Churchland, applies to biological altruism as well.

31. I'll not focus on the vagueness of definitions—let's assume that we figured out the difficulties mentioned on the Definition (D) by providing a precise definition that clarifies all the vagueness mentioned. I believe that this definition is perfectly possible. But even after having such flawless definition, we would still have problems with the concept.

and outcomes of actions; (3) the assumption of the scientific verifiability of one's mental states; and (4) the difficulty of distinguishing altruism and egoism.

The first problem with psychological altruism is the fact that it requires the *actuality* of the mental states as described by Folk Psychology. Psychological altruism only makes sense if we assume some sort of intention or desire underlying actions and a set of beliefs about such actions. Basically, in order to perform an altruistic action, the performer should have an intention/desire to do something that this performer believes would cause a benefit to someone *and* to possess a belief that such action will not bring a benefit for the performer[32]. The assumption of these mental states is problematic. Many philosophers argued against the actuality of Folk Psychology or, at least, against the possibility of using its concepts in scientific research (e.g., Ryle, 1949; Dennett, 1978; Rorty, 1979; Churchland, 1985). Even though I believe that this discussion could lead to a convincing argument against psychological altruism, it would require a long presentation of this complex debate. It will be not necessary here. Instead, I'll focus on other simpler problems.

The second problem of psychological altruism is that the altruist action itself and its outcome are *not* a relevant part of altruism: what really matters are the mental states of the performer. As a case to exemplify how the action is not important to decide if a behavior is a case of altruism, notice that, if a computer is programmed to destruct itself if it put people in danger, no one would take its auto-destruction as an altruist act, even though it acts for other's sake without having any benefit from such action. Why can't the robot be

32. Or, at least, *not have* any belief that states that some benefit would follow from this action.

altruistic? Because the robot doesn't have the appropriate intention or desires motivating its actions.

What is interesting in the example above is that the same would be true for a human being if there was not an intention behind the specific action in question. Consider a second example: there is one piece of poisoned fruit in a basket placed on the table of a crowded room and someone unknowingly eats the poisoned fruit (thus saving the life of someone else who would have eaten it). Why is this action not altruistic? Because, although the action resulted in a loss to the performer and a benefit to others, it lacks the *intention* to do it. As Sober & Wilson (1998, p. 17) say, "[p]hilosophical and psychological discussions of altruism often concentrate so heavily on motives that the actual act of helping is ignored".

If the examples above aren't enough, consider a third one. Someone gives five dollars to a homeless person. This action could be taken as an altruistic action, but someone may doubt it and say that, in fact, the performer felt pleasure by doing such action, and, therefore, it wasn't an altruistic action: the five dollars wouldn't bring more pleasure for the performer if used somewhere else, then, inserting the five dollars in a vending machine or in the homeless person's hat wasn't so different, after all. The point is: the action itself would be useless to decide if it was an altruistic action or a selfish one.

The irrelevance of action for psychological altruism makes clear that simple observation of behaviors is useless as a method to classify who is acting altruistically. This is a huge limitation if we want to use altruism scientifically, since we are talking about something without a physical observable manifestation. But, if we want to use psychological altruism scientifically, we need to assume some sort of verifiability of the mental states involved in it. This is our third problem with psychological altruism:

apparently there are no good methods to access such mental states. The intuitive method that one may suggest is *introspection*. It is widely accepted in science that introspection, alone, is a problematic and biased method.

> Understandably, there has been resistance, among psychologists and neuroscientists, against self-consciously using introspective (first-person) methods in their scientific studies. For it has been thought self-evident that the deliverances of introspection are not intersubjectively accessible, hence verifiable, and what is not intersubjectively verifiable cannot be the subject matter of science because science is in the business of studying objective reality. The objectivity of science consists, at a minimum, in the intersubjective availability of its subject matter, in that no one is epistemically privileged with regard to gathering evidence about the object of the study (Price & Aydede, 2005, p. 244).

Either way, to make clear why introspection is not a good option here, I'll present how it is inefficient even to make the distinction between altruism and egoism—which is the most basic distinction we would expect from a method that supposedly would show us when altruism occurs and when it doesn't. This will be the fourth and last problem.

Going back to the example of giving five dollars to a homeless person, we need to consider that the difficulty of distinguishing between altruistic and egoistic acts exists both for someone who is observing the performer and for the performer herself. It is common, especially after Freud[33], postulating that the intentions/desires underlying the actions of individuals can be unconscious for the individuals themselves. Therefore, performers may seem to be acting altruistically and even believe themselves to be doing so, yet the real motivation can simply be an egoistic desire.

The idea that every action is egoistic, and that in the roots of every apparently altruistic action there is a selfish desire, is what is known as *psychological egoism*.

33. "Freud and his school maintained that the unconscious conceals and systematically distorts mental contents; one of Freud's deepest influences on psychology was to cast doubt on the reliability of introspection" (Sober & Wilson, 1998, p. 253).

Psychological egoism has been defended through empirical research (e.g. Slote, 1964) and through conceptual analysis, as we can see already in Plato's *Meno* (Plato, 2002). Sober (2000, p. 129) explains:

> Psychological egoism is a theory about motivation that claims that all of our ultimate desires are self-directed. Whenever we want others to do well (or ill), we have these other-directed desires only instrumentally; we care about others only because we think that the welfare of others will have ramifications for our own welfare.

If psychological egoism is true, charitable acts should be interpreted as egoistic actions, just like any other behavior. The supporter of psychological egoism will argue that, even if someone sacrifices money in an act of charity, it was the desire to do it, in the first place, that motivated the action[34]—otherwise there would be no reason or motivation for this individual to perform such an action.

Thus, even accepting the existence of the required mental entities from Folk Psychology *and* some sort of access to them, there would still exist a difficulty to distinguishing egoism and altruism. Considering how difficult it is to be sure about when one is acting altruistically or not and the fact that our motivations usually are vague and pluralistic, we might entertain a third option. This other approach denies that "pure altruism" and "pure egoism" are an adequate description of what is happening in the mind. Instead, it suggests what we might call "impure altruism" and "impure egoism" (e.g. Ottoni-Wilhelm, Vesterlund, & Xie, 2017). While pure altruism requires a completely altruistic intention underlying an action, impure altruism accepts some altruistic motivation mixed with some egoistic motivation. The flexibility of altruism, by proposing a weaker

34. To see ourselves as altruistic, for example, may result in a pleasure that overcomes the economic loss in the act of charity. Explanations such as this can be provided to every other altruistic action.

version, helps to dissolve a case of *false dilemma* between altruism and egoism in the theory, but, at the same time, makes a scientific classification of an action in terms of altruism or egoism, that was already difficult, virtually impossible.

Considering psychological altruism's limitations presented here, one may ask: how could a concept as obscure as psychological altruism to be so popular? I'll offer a hypothesis that can answer this question. In short, my hypothesis is that altruism, as it is used in everyday life—I'll call it "common-sense altruism"—, is so widely used because (1) common-sense altruism is not identical to psychological altruism: it is the combination of psychological altruism (because it is about mental states) with a positive moral valence attached to it; and (2) common-sense altruism is not used for *descriptive* purposes, but for *normative* ones.

As an example of a common-sense altruism, consider that we are trying to know if the veganism of a friend is altruistic or nor. Avoiding the consumption of products derived from animals is a behavior likely to be taken as altruistic, for obvious reasons. Also, our friend seems to be really concerned with the well-being of animals. We can say that if there are observable conditions for altruism, they are satisfied. At this point, we may assert that this friend's veganism is, in fact, altruistic. But, at the same time, the opposite thesis, namely, that despite the appearances our friend is not altruistic, remains a possibility, since psychological altruism is about mental states not about observable behaviors. There isn't a clear criterion to make this distinction. Consequently, we can ask: considering the standard way in which we use altruism in the common-sense attributions, how would we decide if our friend's veganism should be taken as altruistic or not?

My hypothesis is that we will accept or deny altruism not trying to be descriptively accurate, but rather trying to dignify what we believe to be morally praiseworthy or to reduce what we believe to be morally unworthy—depending on the practical convenience of our friend's veganism, we will be inclined to take it as altruistic or not.

My hypothesis is an empirically testable, but here I'm only defending it theoretically. My background to proposing such hypothesis is a pragmatic account of how we use language in everyday life. Instead of considering our use of language in everyday life as something that constantly tries to be more objective, I consider the objectivity of our everyday life language as only one of the goals of our complex way of using language. In many situations, the value of objectivity is neither the only value available, nor the most important one.

If we consider that, in everyday life, we use words in the most advantageous way possible, exploring all their capacities, considering, for example, the outcome of our use of a given term, we can postulate that the vagueness of a word will not necessarily be something that we want to avoid. Instead of interpreting the vagueness from common-sense altruism as an unfortunate *limitation*, we would be far more realistic if we take it as a *feature*, which is as fundamental for the usage we give to this term as any other property of it. Supporting what we evaluate morally seems to be a clear advantage, and the imprecision of common-sense altruism is an efficient tool for achieving this goal.

I need to make clear that I'm not suggesting, here, that we should stop addressing the phenomenon of psychological altruism. The psychological processes that result in actions that benefit others should be studied. In fact, Sober & Wilson (1998) provide good arguments for the thesis that humans are biologically hard-wired to act not only to benefit

the performer, neither only to seek pleasure, but to benefit other as well. The point I want to make is that if it is the case that common-sense altruism is used in the way I presented, it is fundamental that we make proper distinctions between it and the scientific concept of psychological altruism, which should be strictly descriptive.

Before finishing this section, I will stress the distinction of each altruism addressed here. We can sum it up in this way: (1) biological altruism is a descriptive term, and is about the observable effects of a given behavior; (2) psychological altruism is a descriptive term, and is about the mental states motivating the individual; and (3) common-sense altruism—which I propose—is the combination of psychological altruism and a moral evaluation of actions, and is used for practical purposes.

In the first chapter I addressed some philosophers trying to use scientific knowledge to justify moral claims. As I concluded, their morality is as scientific as a religious person's morality is divine. In this chapter I focused on the science underlying the discussion started on the first chapter. With the example of Kropotkin, I explored ways in which scientific observation itself can be influenced by non-cognitive values and how vagueness is a problem for science. In this last section, with the study case of altruism, I tried to show the role of scientific terminology in this discussion, explaining how it can benefit our philosophical debate (e.g., excluding vagueness) and how it can also inhibit it (e.g., making too narrow definitions, which make the term no longer useful for the debates in which we used it).

In any case, the way in which philosophy and science overlap in the problems addressed here hopefully will highlight the importance of interdisciplinarity. Every philosopher will gladly point to the dangers of doing science while ignoring philosophy.

On the other side, every scientist, in the same way, will gladly point to the dangers of doing philosophy while ignoring science. This thesis indicates that both are right. Blindly adopting scientific terminology without criticizing it and understanding its limitations is a mistake. But ignoring science, as many philosophers proudly do, using terminology from everyday life, claiming that it can shows us something deeper than what science can reach, makes the philosopher vulnerable to the tides of irrationality that permeate the way in which we humans use this tool called language.

# References


Ayer, A. J. (Ed.). (1966). *Logical Positivism*. Simon and Schuster.

Barragan, R. C., & Dweck, C. S. (2014). Rethinking Natural Altruism: Simple Reciprocal Interactions Trigger Children's Benevolence. *Proceedings of the National Academy of Sciences, 111*(48), 17071-17074.

Cabanac, M., Cabanac, A. J., & Parent, A. (2009). The Emergence of Consciousness in Phylogeny. *Behavioural brain research*, 198(2), 267-272.

Campbell, R. L. (2006). Reply to Robert H. Bass, "Egoism versus Rights" (Spring 2006): Altruism in Auguste Comte and Ayn Rand. *The Journal of Ayn Rand Studies*, 357-369.

Carnap, R. ([1950] 1962). *Logical foundations of probability*. The University of Chicago Press.

Carnap, R. ([1958] 2012). Edited by Martin Gardner. *An Introduction to the Philosophy of Science*. Courier Corporation.

Carruthers, P. (2008). Meta-cognition in Animals: A Skeptical Look. *Mind & Language*, 23(1), 58-89.

Churchland, P. M. (1985). Reduction, Qualia, and the Direct Introspection of Brain States. *The Journal of Philosophy, 82*(1), 8-28.

Churchland, P. S. (2011). *Braintrust: What Neuroscience Tells Us About Morality*. Princeton University Press.

Clavien, C., & Chapuisat, M. (2013). Altruism Across Disciplines: One Word, Multiple Meanings. *Biology & Philosophy*, 28(1), 125-140.

Comte, A. ([1842] 2000). *The Positive Philosophy of Auguste Comte.* Translated by H. Martineau. (Vol. 1). Kitchener: Batoche Books.

Comte, A. ([1848] 2009). *A General View of Positivism*. Cambridge University Press.

Comte, A. (1877). *System of Positive Polity: Social Dynamics; or, the General Theory of Human Progress* (Vol. 4). Longmans, Green and Company.

Comte, A. ([1891] 2009). *The Catechism of Positive Religion: Or Summary Exposition of the Universal Religion in Thirteen Systematic Conversations Between a Woman and a Priest of Humanity*. Cambridge University Press.

Darwin, C. ([1859] 2001). *On the Origin of Species*. The Pennsylvania State University Press.

Dennett, D. C. (1978). Why you Can't Make a Computer that Feels Pain. *Synthese, 38*(3), 415-456.

Dubs, H. H. (1951). The Development of Altruism in Confucianism. *Philosophy East and West*, 1(1), 48-55.

Dugatkin, L. A. (2011). *The Altruism Equation: Seven Scientists Search for the Origins of Goodness*. Princeton University Press.

Dugatkin, L. A. (2007). Inclusive Fitness Theory from Darwin to Hamilton. *Genetics*, 176(3), 1375-1380.

Emerson, R. W. (1940). *The Complete Essays and Other Writing.* Ed. Brooks Atkinson. New York: Modern Library.

Haber, M. (2013). Colonies are Individuals: Revisiting the Superorganism Revival. *From Groups to Individuals: Evolution and Emerging Individuality,* 195-217.

Hamilton, W. D. (1964). The Genetical Theory of Social Behavior I and II. *J. theor. Biol*., 7, 1-52.

Hare, R. M. ([1952] 1991). *The Language of Morals*. Oxford University Press.

Harman, O. (2014). A History of the Altruism–morality Debate in Biology. *Behaviour*, 151(2-3), 147-165. Chicago.

Harris, S. (2004). *The End of Faith, Religion, Terror and The Future of Reason*. N. Y.: Norton.

Harris, S. (2010). *The Moral Landscape: How Science Can Determine Human Values*. Simon and Schuster.

Hempel, C. G. (1966). *Philosophy of Natural Science*. Prentice Hall, Inc.

Hempel, C. G. ([1950] 1966). The Empiricist Criterion of Meaning. In Ayer, A. J. (Ed.). *Logical positivism,* 108-132. Simon and Schuster.

Hume, D. ([1738] 2003). *A Treatise of Human Nature*. Courier Corporation.

Huxley, T. H. (1902). *Evolution and Ethics: And Other Essays* (Vol. 9). D. Appleton.

Kropotkin, P. A. ([1902] 2012). *Mutual Aid: A Factor of Evolution*. Courier Corporation.

Kropotkin, P. A. (1970). *Kropotkin's Revolutionary Pamphlets: A Collection of Writings*. Dover Publications.

Kropotkin, P., & Walter, N. (Ed.). (2010). *Memoirs of a Revolutionist*. Courier Corporation.

Lacey, H. (2005). *Is Science Value Free?: Values and Scientific Understanding*. Taylor & Francis e-Library.

Leinfellner, W. (1985). A Reconstruction of Schlick's Psycho-sociological Ethics. *Synthese*, *64*(3), 317-349.

Lewisohn, D. (1972). Mill and Comte on the Methods of Social Science. *Journal of the History of Ideas*, *33*(2), 315-324.

MacIntyre, A. C. (1969). Hume on 'is' and 'ought'. In *The Is-Ought Question*, 35-50. Palgrave Macmillan, London.

Mech, L. D., & Boitani, L. (Eds.). (2010). *Wolves: behavior, ecology, and conservation*. University of Chicago Press.

Moore, G. E. ([1903] 2012). *Principia Ethica*. Courier Corporation.

Nagel, Thomas. (1978). *The Possibility of Altruism*. Princeton University Press.

Ottoni-Wilhelm, M., Vesterlund, L., & Xie, H. (2017). Why Do People Give? Testing Pure and Impure Altruism. *American Economic Review*, 107(11), 3617-33.

Plato (2002). *Meno*. Trans. by J. Holbo & B. Waring.

Poincaré, H. ([1913] 1963). Ethics and Science. In *Mathematics and Science: Last essays*. Translated by John Bolduc. Dover Publications, Inc., New York.

Poincaré, H. (1907). *The Value of Science*. Translated by George Bruce Halsted. Science Press.

Price, D. D., & Aydede, M. (2005). 13 The Experimental Use of Introspection in the Scientific Study of Pain and Its Integration with Third-Person Methodologies: The

Experiential-Phenomenological Approach. Ed. Murat Aydede. In *Pain: New Essays on Its Nature and the Methodology of Its Study*, 243-273.

Rand, A. (1964). *The Virtue of Selfishness*. Penguin.

Renwick Monroe, K., Martin, A., & Ghosh, P. (2009). Politics and an Innate Moral Sense: Scientific Evidence for an Old Theory?. *Political Research Quarterly*, 62(3), 614-634.

Rhode, D. L. (2005). *Pro Bono in Principle and in Practice: Public Service and the Professions*. Stanford University Press.

Rorty, R. (1979). *Philosophy and the Mirror of Nature*. New Jersey: Princeton University Press.

Ryle, G. ([1949] 2009). *The Concept of Mind*. Routledge.

Schlick, M. ([1930] 1966). The Turning Point in Philosophy. In Ayer, A. J. (Ed.). *Logical Positivism, 53-39*. Simon and Schuster.

Schlick, M. (1939). *Problems of Ethics*. Trans. David Rynin. New York.

Schlick, M. ([1948] 1966). Positivism and Realism. In Ayer, A. J. (Ed.). *Logical Positivism, 82-107*. Simon and Schuster.

Sharma, R. K. (1996). *Fundamentals of Sociology*. Atlantic Publishers & Dist.

Slote, M. A. (1964). An Empirical Basis for Psychological Egoism. The Journal of Philosophy, 61(18), 530-537.

Sober, E. (2000). Psychological Egoism. *The Blackwell Guide to Ethical Theory*. Blackwell, Malden, MA, 129-148.

Sober, E., & Wilson, D. S. (1998). *Unto Others: The Evolution and Psychology of Unselfish Behavior*. Harvard University Press.

Stich, S. (2016). Why There Might Not Be an Evolutionary Explanation for Psychological Altruism. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences, 56*, 3-6.

Stroebe, W., Postmes, T., & Spears, R. (2012). Scientific Misconduct and the Myth of Self-correction in Science. *Perspectives on Psychological Science*, 7(6), 670-688.

Todes, D. (1987). Darwin's Malthusian Metaphor and Russian Evolutionary Thought, 1859-1917. *Isis*, 78(4), 537-551. Retrieved from http://0-www.jstor.org.carlson.utoledo.edu/stable/231917

Todes, D. (1989). *Darwin Without Malthus: the Struggle for Existence in Russian Evolutionary Thought*. Oxford University Press on Demand.

Velleman, J. D. (1988). Brandt's Definition of "Good". *The Philosophical Review*, 97(3), 353-371.

Wernick, A. (2001). *Auguste Comte and the Religion of Humanity: the Post-theistic Program of French Social Theory*. Cambridge University Press.

Wilson, D. S. (2015). *Does Altruism Exist?: Culture, Genes, and the Welfare of Others*. Yale University Press.

West, S. A., Griffin, A. S., & Gardner, A. (2007). Social Semantics: Altruism, Cooperation, Mutualism, Strong Reciprocity and Group Selection. *Journal of Evolutionary Biology*, *20*(2), 415-432.