

THE ETHICS AND SOCIAL PHILOSOPHY OF DAVID K. LEWIS'S MODAL REALISM

Written by Panagiota Nigianni
Under the supervisory guidance of
dr. Paul Dekker

Masters Thesis (18 ECTS)
Degree of Master of Arts (MA)

Philosophy

Department of Philosophy
Graduate School of Humanities



UNIVERSITY OF AMSTERDAM

Acknowledgments

I thank the philosophy tutors on the MA and (r)MA Philosophy programs, also those on the MSc Logic program, at UvA, for taking on the difficult task of teaching and promoting research across methodologies and areas of philosophy. I would especially like to thank tutors teaching philosophy of language, ethics and social and political philosophy, who have been influential, if only indirectly, on the conception, articulation and advancement of the main thesis and arguments of my study of Lewis's ethics and social philosophy. I was fortunate to receive early tutorial advice on possible worlds, modal metaphysics and Lewisian 'genuine' modal realism by prof. Francesco (Franz) Berto. I would also like to thank the Graduate School of Humanities (GSH) at UvA, for supporting new ideas and facilitating innovative research in philosophy. Finally, special thanks to my supervisor, dr. Paul Dekker at GSH/ILLC UvA, for his rigorous supervision, but also for his cooperative attitude, for keeping an open mind and persevering in support of my work.

Contents

1.	Introduction: Research objective	3
1.1.	Summary	3
1.2.	Overview of theoretical context and research subject	4
2.	Lewisian 'genuine' modal realism	8
3.	Lewis's philosophical method of analysis	16
3.1.	The example of Lewis's dispositional theory of value	20
4.	Lewis's value theory: Dispositional theory of value as a solution to the problem of the anti-Humean desire as belief thesis	22
5.	Lewis's folk 'moral science'	30
5.1.	Lewis's conventionalist approach	33
5.2.	Lewis's modal metaphysical approach	38
6.	Lewis's value pluralism	43
6.1.	Lewis's deontic ethics in sphere semantics	52
7.	Lewis's contributions to metaethics, ethics, social and political philosophy	56
7.1.	Overview of Lewis's publications in ethics and social philosophy	56
7.2.	Metaethics	58
7.3.	Ethics, social and political philosophy	73
8.	Challenging the canon: Lewis's positioning in respect to canonical accounts	82
9.	Conclusion	99
	Bibliography	102
	Appendixes	III
I.	Objections and replies	112
II.	Glossary	126
III.	Index of Lewis's symbolic notation	134

I

Introduction: Research Objective

I.1 Summary

The research subject of this thesis is the ethics and social philosophy of the analytic philosopher David Kellogg Lewis. In this research I set out to answer the question of what Lewis has contributed to canonical accounts of ethics and social and political philosophy. The evident research requirement of situating and evaluating Lewis's ethics and social philosophy in respect to systematic canonical accounts derives from Lewis's inconclusive and incomplete academic effort to systematize this aspect of his work. At the same time, pursuing this research objective serves the purpose of making the select topic a tangible and accessible object of rigorous philosophical study. For this purpose, I examine Lewis's interpretations of moral philosophical problems in his essays in ethics and social philosophy in view of Lewis's modal realist philosophy.

Undertaking the research task of bringing together, presenting and evaluating Lewis's contributions in ethics and social philosophy, I have come across the analysis, indispensable to this research, of Lewis's arguments about his defense of the Humean against the anti-Humean thesis, which are formulated in a schematic, yet systematic, manner in Lewis's value theory. I show that Lewis builds upon this critique, whilst drawing upon his own philosophico-logical project of 'genuine' modal realism, by applying systematic thinking in his accounts of folk 'moral science' and of possible worlds pluralism, where Lewis advances his metaphysical view of value pluralism. I conclude that, in counter arguing principles of select canonical accounts and offering counterexamples to normative ethical problems, Lewis's ethics and social philosophy debate the canon, by deriving solutions from Lewis's value theory and its applications in folk theory and possible worlds value pluralism.

1.2 Overview of theoretical context and research subject

Lewis's impressive and innovative philosophical project has primarily been known for its revalidation of metaphysics, which came as a reaction to positivist and neo-positivist analytic philosophy in the second half of the twentieth century.¹ In modern philosophy, empiricist and pragmatist thought have had a negative influence on the philosophical realm of metaphysics. Considering Quinean naturalism, there is an obvious, more or less, complete rejection of metaphysics, especially modal metaphysics, which is the philosophical strand where Lewis's project has been situated.² Traditionally, the logical positivists have argued that metaphysical statements are meaningless, because they do not make any predictions about possible experience. As van Inwagen and Zimmerman (2008: 6) point out, the logical positivist thesis claims that: "The meaning of a statement consists entirely in the predictions it makes about possible experience".

Characterized by methodological systematization, the revalidation of analytic metaphysics addressed an early and renewed interest towards, amongst others, ontological concerns (Quine, 1969, 1976; Prior and Fine, 1977) by means of the logic of modality.³ This renewed interest places Lewis's philosophy, especially his metaphysics and his philosophico-logical project of modal realism, in relation to the broader developments in modern and contemporary analytic philosophy.

¹ In this context, Lewis has been characterized as one of the greatest metaphysicians and philosophers at the turn of the twenty first century. Divers (2007: 74) makes the claim that Lewis and

² According to Divers (2007: 71-72), early Wittgensteinian explicit substantive metaphysics of logical modality and Carnapian intensional metaphysics of linguistic modality have provided an exception by maintaining an interest in modal metaphysics. Quine (1969: 91) himself problematizes the metaphysical question of existence in light of the logical positivist view, specifically the Carnapian view, of whether the questions "whether there are numbers, or qualities, or classes", or whether there are bodies, are metaphysical questions, thus meaningless; while questions such as whether there are "prime numbers between 10 and 20", or whether "there are rabbits, or unicorns", are non-metaphysical, thus meaningful. Quine (1969: 91-92) explains that whenever there was use of category words, Rudolf Carnap ruled meaningless questions of existence. However, Quine (1969: 92) explains that category comprises of "the range of some distinctive style of variables", where the style of variable is understood to be arbitrary, thus not helpful in distinguishing between meaningful and metaphysical questions of existence. On the other hand, Lewis (1986a: 141) characterizes both Carnap and Quine as linguistic ersatzists; although Lewis (1986a: 145) specifically recognizes that Carnap's "worldmaking languages" are presented with interpretation, to construct "state-descriptions or models" (Lewis, 1986a: 152).

³ Regarding Quine's view of modal realism, Lewis (1986a: 217) argues that he "portrays a form of modal realism that treats ordinary things as transworld individuals, perduring through non-overlapping worlds in just the way they perdure through time and space". Garson (2006: 425) puts forth a defence against Quine's objections towards quantifying into, so-called, "opaque contexts", partly because of his rejection of essentialism, by pointing that "even if sentences that quantify in make assertions that are philosophically objectionable, this is hardly a reason to ban quantifying in from logic" (Garson: 2006: 430).

According to Burge's critical historical evaluation of logical positivism (1992: 27-28), the movement's decline due to factors internal to the subject of the philosophy of language⁴ signalled the dismissal of the philosophy of language as the dominant starting point for philosophical activity in Anglo-American analytic philosophy from the late 1970's onward (Nolan, 2005: 3). Instead, there has been a demarcated shift of starting points to other philosophical areas, including the philosophy of mind (Burge, 1992: 28). Consistently with this shift, as I discuss in **section 3**, Lewis (1983c: xi) objects to an analysis of thought and modality, which sets off solely from language itself, without bypassing the significance of language for analytic philosophy,⁵ by proposing instead to reconsider language in terms of semantics integrated more broadly into language's use in social interaction. This Lewisian position also informs Lewis's philosophical method of analysis, which I present in **section 3**, and Lewis's folk theory, which I present in **section 4**.

Although Lewis's philosophy systematized and built upon the study of different areas, including metaphysics, philosophical logic, philosophy of language, philosophy of mind, and value theory (Hall, 2010), different parts contain significant fragments of his work on ethics and social philosophy. In his monograph *David Lewis* (2005), Nolan (2005: 178) argues that this lateral aspect of Lewis's philosophy is not as well known. In my assessment, which also motivates the research objective of this thesis, Lewis's work on ethics and social philosophy, which is an exception for analytic philosophy, has been undervalued and insufficiently researched, because neither has it been adequately contextualized in any area of canonical ethics and social and political philosophy, nor has it been considered systematic.⁶ As Lewis (2005: 320) himself modestly states, he has only offered a fragment of a moral system in his value theory, which is explained in his essay "Dispositional Theories of Value" (1989). Furthermore, besides presenting a rather uncharted territory for philosophical study, this aspect of Lewis's work bears obvious challenges of a different sort, namely methodological, interpretative and

⁴ Burge (1992: 4) comments on the logical positivists' view of the verificationist principle:

"The verificationist principle was supposed to explain why philosophy, particularly metaphysics, had failed. The idea was that since philosophy associates no method of verification with most of its claims, those claims are meaningless. To be meaningful and produce knowledge, philosophy was supposed to imitate science in associating its claims with methods of testing them for truth".

⁵ Therefore, as Nolan (2005:3) explains, "in some broad sense and other traditions (not only continental)", Lewis was an analytic philosopher.

⁶ Nolan (2005: 186) argues that Lewis never advances a complete ethical system, while he could have thought that ethics may not be systematic.

argumentative, in the context of analytic philosophy. As Priest, cited in Quinn's review "Papers in Ethics and Social Philosophy by David Lewis" (2004), eloquently states:

"Lewis works like this: he gets interested in puzzles and problems; he likes to solve them; he does so by applying his technical expertise, his great ingenuity, his prowess in the thrust, party, and counter-thrust of philosophical debate." (Quinn, 2004: 712)

In response, I show in **section 4** that Lewis offers a more or less systematic value theory account derived from his critique of the anti-Humean thesis. I also show that, building upon this critique, but also drawing from his philosophico-logical project of 'genuine' modal realism, Lewis applies systematic thinking in his accounts of folk 'moral science' and of possible worlds pluralism, as I explain in **sections 5 and 6**. Finally, I show in **section 7** that Lewis's contributions engage with a range of canonical systematic discourses in metaethics, such as: determinism and non-determinism; compatibilism and incompatibilism; conventionalism; realism, quasi-realism and fictionalism; while, in his social and political philosophy, Lewis addresses utilitarianism, which is a strand of consequentialism. In view of **section 7**, in **section 8**, I evaluate Lewis's overall contributions to ethics and social philosophy by showing that Lewis's work debates canonical accounts. Besides offering a critical positioning of Lewis's ethics and social philosophy in respect to canonical accounts, in this last section I place emphasis on Lewis's modal realist interpretations of canonical views, in consistence with Lewis's philosophy of modal realism.

On an additional note, the demanding complexity, breadth and variety of the Lewisian philosophical project presupposes that one has a minimal understanding of Lewis's philosophical commitments and methods of working, before delving into a specific area or aspect of the philosopher's work, such as his value theory and his ethics and social philosophy. For this reason, I have included in the **Appendixes**, besides a glossary of standard terms used by Lewis, also core problems of Lewis's philosophy discussed in the form of 'objections' and 'replies',⁷ which are complementary to topics covered in the main thesis. In this way, I aim to offer guidance and additional insight to the lay reader with an interest in Lewis.

In view of the coherence and systematicity of Lewis's philosophy, Hall (2010) suggests that, one might think that Lewis would recommend doing philosophy in a

⁷ This is a form of argumentation, which Lewis (1983e [1975a]) himself uses in counter arguing possible objections to his claims.

specifically ordered manner, which is a potential misunderstanding. For Hall (2010), Lewis rather recommends a "holistic approach" to philosophy:

"We start with a total body of claims we are inclined to believe – whether on the basis of "common sense" (an oft-invoked category, for Lewis) or of science – and try our best to *systematize* it in accordance with standards of theoretical goodness that are themselves endorsed by common sense and/or science (and so are themselves, to some extent, also up for grabs). A substantial portion of Lewis's overall body of philosophical work can thus be seen as an extended – and breathtakingly ambitious attempt at achieving *total reflective equilibrium*."

Hall (2010) here explains Lewis's claim (1973: 88) that, despite one coming to philosophy already holding "a stock of opinions", "it is not the business of philosophy either to undermine or to justify these preexisting opinions, to any great extent, but only to try to discover ways of expanding them into an orderly system". With this thought, I now move on to present Lewis's philosophico-logical project of 'genuine' modal realism.

2

Lewisian 'genuine' modal realism

As I argue in the forthcoming sections, the philosophico-logical project of modal realism pertains to Lewis's work on ethics and social philosophy. Here, I offer a brief exposition of Lewis's own version of modal realism, which is Lewis's philosophical system derived from the logical notion of possible worlds, which is also a metaphysical systematic notion in Lewis's philosophy. As Lewis (1983c: ix), also cited in Nolan (2005: 1), has famously argued: "I should have liked to be a piecemeal, unsystematic philosopher, offering independent proposals on a variety of topics. It was not to be".

Broadly speaking, much of the discussion on modal realism has developed on the possible worlds notion, which is a modal primitive (Brandom, 2014: 147), allowing for modal reasoning, but also, broadly, for answering ontological questions of realism and anti-realism. Brandom (2014: 198-199) has characteristically stated that the innovation of possible worlds semantics was a "modal revolution" in twentieth century Anglophone philosophy, starting with Kripke's algebraic possible worlds semantics and continuing with the generalization of Kripke's "apparatus to an intensional semantics for non-logical expressions" by Lewis and others, such as Stalnaker and Montague. As Burge (1992: 48) argues, thought experiments have been used in philosophical analysis to give "new forms to many old issues"; in this view, the possible worlds hypothesis can also be understood as a thought experiment for contemporary metaphysics. Unarguably, possible worlds have advanced the validity of modal reasoning in philosophical logic, the philosophy of language and metaphysics. For Lewis (1986b: 9), modal reasoning has been clarified by "a semantic analysis of modal logic by reference to possible

worlds and to possible things therein".⁸ As Lewis (1986b: 3) explains, "our understanding of modality has been much improved by means of possible worlds semantics, which is the project of analyzing modal language by systematically specifying the conditions under which a modal sentence is true at a possible world".⁹

Simply put, Lewisian modal realism is the philosophical position that there are possible worlds, other than the one we inhabit, which are just as real or concrete as ours without lacking specificity (Buehler, 2014: 25; Lewis, 1986a: 86)¹⁰ and without being abstracted from something, such as an entity, or a set of entities. This definition is consistent with Lewis's argument from natural language (Buehler, 2014: 25), according to which: first, it is uncontroversial that things could have been different in many ways than the way they are, which Lewis (1986a: 86) calls "ways a world could possibly be", while he assigns this definition to worlds themselves; secondly, this is a statement to be taken at face value, as a "serviceable hypothesis" of which we have no reason to think it is not true (Lewis, 1986a: 3).¹¹ In Lewis's own words (1986a: viii):

"I must insist that my modal realism is simply the thesis that there are other worlds, and individuals inhabiting these worlds; and that these are of a certain nature, and suited to play certain theoretical roles."

⁸ Besides the Lewisian claim cited above, Kripke (1980: 44) has argued that we should not think of possible worlds as things, which we can learn anything about by observation; as, for instance, in his claim: "A possible world isn't a distant country that we are coming across, or viewing through a telescope". Instead, we should think of them in epistemological terms, an approach "determined by the descriptive terms we associate with them" (Gamut, 1991: 71). Despite their mind-independent ontological existence, for Kripke, as Gamut (1991: 71) argues, possible worlds are not things for us to discover; instead they need to be introduced in terms of being stipulated. Notably, Gamut (1991: 72) mentions that the Kripkean epistemic interpretation is comparatively contrasted with the Lewisian ontological and 'extremely' realist conception of possible worlds (1986a: 104), which is also mind-independent, thus within the realist philosophical tradition.

⁹ Objections against possible worlds, as a purely metaphysical notion – beyond its contemporary algebraic conception – lacking empirical content, have also been raised by means of objections to intensional logic. Gamut (1991: 71) argues that:

"Everything which is analysed in terms of it - the concept of intensionality, the concepts of necessity and possibility, the modalities *de dicto* and *de re*, and many more – consequently remain as obscure as they were in the first place. Although it seems as though they have been clarified, it is argued, intensional logic really only succeeds in substituting one murky notion for another".

¹⁰ Although worlds are concrete in Lewisian modal realist terms, regarding parts of worlds, Lewis (1986a: 86) explains that some of them are concrete, but perhaps not all: for instance, in the case of universals, or particulars that are non-spatiotemporal parts of ordinary particulars, which are parts of worlds, Lewis (1986a: 86) accepts that we have abstractions, which are also parts of worlds.

¹¹ As Buehler (2014: 25) mentions, Lewis's argument from natural language supports the view that we should not take every statement at face value; however, "there is a presumption to do so unless it causes trouble and there is an alternative which does not". There has been no satisfactory argument why modal realism is troublesome and that there is a non-troublesome alternative. As Lewis (1983c: x) argued: "Philosophical theories are never refuted conclusively. [...] The theory survives its refutation – at a price. [...] what we accomplish in philosophical argument: we measure the price."

For Lewis (1986a: 59), the theoretical model of possible worlds is not doctrine, but useful terminology, which combines "set theory and modal realism in structured and unstructured versions of properties, relations and propositions". Lewis (1986a: 234) is aware that new terminology does not make up for a new theory, but rather works on a modification of a previous one. Despite Lewis's recognition (1986a: 136-137) that his version of modal realism disagrees to a certain degree with common sense, which may be more compatible with actualism, the view that there is one actual world, ours, while other worlds may exist as abstract representations, Lewis (1986a: 134) explains that, with his modal realism, he aims to improve the "unity and economy" of the total theory of modality. Lewis dares further to characterize modal realism as an improvement in "unity and economy" over our common sense opinions, because the advantages of modal realism outweigh the costs of disagreeing with common sense (Nolan, 2005: 57).

In Lewisian modal realism (1986a: 187), propositions are sets of worlds,¹² which cannot be considered as abstractions, unless they are equivalence classes "only under thoroughly artificial equivalences" (Lewis, 1986a: 85). Lewis (1986a: 85) argues:

"I emphatically do not identify possible worlds in any way with respectable linguistic entities; I take them to be respectable entities in their own right. When I profess realism about possible worlds, I mean to be taken literally. Possible worlds are what they are, and not some other thing. If asked what sort of thing they are, I cannot give the kind of reply my questioner probably expects: that is a proposal to reduce possible worlds to something else."

Characteristically, in *The Plurality of Worlds* (1986), Lewis's metaphysics makes use of the modal notion of *possibilia*. By using this notion in the possible worlds hypothesis, Lewis advances his earlier work on universals and particulars,¹³ but also

¹² Other than Lewis's version of 'genuine' modal realism, there are modal realist strands, which Lewis defines as 'ersatzist', for which possible worlds are simply sets of 'ersatz' worlds. Lewis (1986a: 185) emphatically states that he is not opposed to "states of affairs, ways things might be, possibilities, propositions or structures"; still, for Lewis, this means that he believes in entities that are suited to play the roles, which their names ascribe to them. His proposed entities are sets of worlds and are the same in every case: "Worlds as I understand them: us and all our surroundings, and other things like that" (Lewis, 1986a: 185); see also **Appendix I**, "Objections and Replies", 3 and 4.

¹³ In his earlier work on metaphysics, Lewis (1983b: 344-345) pertains to his distinction between universals and properties, accepting that universals are essentially intrinsic and constituent parts of particulars. Universals are sparse and "should comprise a minimal basis for characterizing the world completely" (Lewis, 1983b: 345-346); unlike particulars or properties, which are conceived as abundant, or, sometimes, as sparse, dependent on whether they are extrinsic, or intrinsic, respectively (Lewis, 1986a: 59-60). In contrast with particulars, which follow intuitive principles, universals are not intuitive, because they occur repeatedly, allowing for co-presence, thus capturing facts of objective resemblances and accounting for the causal laws of things (Lewis, 1983b: 345-346). Contrary to universals, properties are understood as classes or as sets, therefore they are

on what he calls 'possibles', which is another modal notion introduced in Lewis's *Convention: A Philosophical Study* (1969). Lewisian modal realism in possible worlds relies on a set theoretical construction of Lewis's metaphysical system of universals and properties for integrating his modal notion of *possibilia*. Lewis takes possible worlds in canonical terms, meaning taking a possible world as a 'way' things might have been in a 'total way';¹⁴ as Lewis (1973: 84) states:

"Ordinary language permits the paraphrase: there are many ways things could have been besides the way they actually are. On the face of it, this sentence is an existential quantification. It says that there exist many entities of a certain description, to wit 'way things could have been'. I believe that things could have been different in countless ways; I believe permissible paraphrases of what I believe; taking the paraphrase at its face value, I therefore believe in the existence of entities that might be called 'way things could have been'. I prefer to call them 'possible worlds'."

In Lewisian modal realism, set-theoretic constructions out of *possibilia* can serve for propositions (Lewis, 1986a: 57). Lewis (1986a: 42) offers the typical example of different individual speakers at different worlds, whose speaker-relativity requires that *possibilia* enter the picture, no matter what the semantic values are themselves

determinant of members of a class; in the Lewisian ontology (1983b: 344), they are taken to be classes of *possibilia*. Universals consist of their simpler constituents, thus they are not distinct entities (Lewis, 1983b: 346), contrary to properties, which have independent ontologies to the things that belong to them; therefore, they do not belong *per se* to certain particulars (Lewis, 1983b: 344-346). Advancing Quine's account (1969), Lewis (1983b: 351) argues that properties enable us to define the content of our intentional attitudes, thus, there is no essentialism attached to them, but rather they are defined by what we refer to. This is a crucial argument for Lewis's notion of belief-desire introduced in his value theory. For the non-essentialist treatment of identity in possible worlds, see also (Chisholm, 1967).

With the introduction of *possibilia*, we also introduce sets of *possibilia*, which are modal entities taking over one aspect of what properties do in the Lewisian system. In view of the above, Lewis (1986a: 85) clarifies in his later work on the plurality of worlds that universals are not abstractions, like names, spatiotemporal locations, causal networks, or body of theory, which cannot be identified with universals or tropes; universals do not constitute part, for instance, of the intrinsic nature of the thing when they are abstracted, but are rather extrinsic.

¹⁴ A canonical understanding of possible worlds provides the basis for modal metaphysics. Plantinga (2003: 103) mentions that the canonical view is characterized by a series of philosophical attempts, including Lewis's early work on possible worlds semantics (Lewis, 1972), to offer a semantical understanding of modal logic based on modal fragments of natural language. Plantinga (2003: 103) offers the following description of the canonical conception of possible worlds:

"Possible worlds themselves are 'taken as primitive', as the saying goes: but by way of informal explanation it may be said that a possible world is a way things could have been – a total way. Among these ways things could have been there is one – call it 'a' – that has the distinction of being actual; this is the way things actually are. 'a' is the one possible world that obtains or is actual; the rest are merely possible. Associated with each possible world W, furthermore, is a set of individuals or objects: the domain of W, which we may call y(W). The members of y(W) are the objects that exist in W; and of course different objects may exist in different worlds".

For the differences in representation between Lewisian 'genuine' modal realism and other modal realists, see **Appendix I**, "Objections and Replies", 3.

and how they are ascribed to the speakers. Lewis (1986a: 42) further argues that *possibilia* may enter into the construction of the semantic values themselves, allowing for modal *possibilist* interpretations, instead of taking truth value as truth in all possible worlds.

The aforementioned move brings forth Lewis's additional innovation in the use of counterfactual conditionals, which was established by the canonical understanding of modality in possible worlds semantics. Lewis (1986b: 3) describes the canonical view as "the project of analyzing modal language by systematically specifying the conditions under which a modal sentence is true at a possible world". Instead of using a strict conditional in modal sentences, Lewis (1986b: 5) uses a variably strict conditional, which allows for assessing the truth of a counterfactual, by ascertaining whether it holds true at some and not all antecedent-worlds:

"Given a far-fetched antecedent, we look perforce at antecedent-worlds remote from actuality. There are no others to look at. But given a less far-fetched antecedent, we can afford to be more fastidious and ignore the very same worlds."

Lewis (1986b: 5) further employs the notion of similarity of worlds to account for his use of a variably strict conditional in possible worlds semantics, arguing that it is not possible to have two worlds, in which the truth of the antecedent is different, while "everything else is just as it actually is": for instance, a world, which is only different from ours, in that Aristotle was not a philosopher. As Lewis (1986b: 5) points out in a similar example of a historical figure:

"Are his predicaments and ambitions there just as they actually are? The regularities of his character?"

Hence, only antecedent-worlds, which are closer to our world in terms of similarity, must be considered for the assessment of counterfactuals.

The Lewisian analysis of counterfactuals is consistent with Lewis's rejection of overlap of particulars in possible worlds; as Lewis (1986a: 205) mentions, "whatever the universals may do, no two worlds have any particular as a common part". Since Lewis (1986a: 53) identifies propositions with certain properties, which are instantiated by entire possible worlds, and properties are sets of their instances, "a proposition is a set of possible worlds". Therefore, in Lewis's view (1986a: 53), propositions *hold* at a world, or are *true* at a world, while other relative holding of propositions requires a shift in the meaning itself of what consists in a proposition

(Lewis, 1986a: 54). Lewis demonstrates that in his version of modal realism, at possible worlds, properties are not instantiations relative to this or that, because such relativism conflates the distinction between relations and genuine properties, which Lewis (1986a: 53) is not prepared to dismiss.¹⁵ Instead, he treats genuine properties as having or lacking *simpliciter*, in parallel to his treatment of relations (Lewis, 1986a: 53). For example, Lewis (1986a: 52) argues that thirst is not a genuine property, since it is temporally dependent: one has it at some times or lacks it at other times. In Lewisian modal realism, one has temporal parts, which are thirsty, in order to be thirsty at various times, thus thirst is not a relational property.

Likewise, for Lewis (1986a: 62), relations of spatiotemporal distance are not internal, because they do not supervene on the natures of their *relata*, which are the structural properties of things, or parts of things relating to each other. Conversely, relations of distance do supervene on the intrinsic nature of the composite of the *relata* taken together. Hence, relations of similarity or difference in intrinsic respects, such as relations of closeness of worlds in similarity as formulated in Lewis's counterfactuals, are internal, while relations of distance are external (Lewis, 1986a: 62). This latter view accounts for the Lewisian modal realist view that possible worlds are causally and spatio-temporally isolated, as well as for Lewis's rejection of transworld causation. For Lewis (1986a: 208), this means that there is no overlap between worlds, in terms of spatiotemporal relations existing between parts of different worlds; nonetheless the parts of a world stand in suitable external relations, preferably spatiotemporal.

Lewis (1986a: 2) goes to the length of identifying the thesis of the plurality of worlds with modal realism itself, in his defense of the position that "a thesis of

¹⁵ An example of Lewis's objection to the notion of the instantiation of properties is Lewis's response to Kripke's modal realism, where Lewis (1986a: 81) problematizes the debated notions of concrete and abstract in the context of possible worlds; for instance, in Kripke's claim of his own objection towards abstraction (1980: 44): "A possible world isn't a distant country that we are coming across, or viewing through a telescope". For Lewis (1986a: 80), this statement is not much of an objection, because the Kripkean view (1980: 44) also follows an abstract conception of worlds:

"A possible world is given by the descriptive conditions we associate with it [...] 'Possible worlds' are stipulated, not discovered by powerful telescopes".

For Kripke (1980: 44), our abstract conception of worlds is not based on their closeness to ours, but on our presumed mode of accessibility to these worlds by so-called "telescoping viewing", a method of gathering information based on a causal process; as Lewis (1986a: 80) mentions, "a 'telescope' which produced images that were actually independent of the condition of the thing 'viewed' would be a bogus telescope". For Lewis (1986a: 80), if there is no transworld causation, there is not a transworld telescope, like there is no transworld travel. Here, the Lewisian approach to causation counter argues a consequentialist inference to abstraction, because the images we receive telescopically are not causally connected with the condition of the viewed world, which they represent.

plurality of worlds, or modal realism" holds that our world is but one among many. Lewis (1986a: 2) is definite in his description of the plurality of worlds as modal pluralism, however he states that this is not a conclusive hypothesis, but rather a useful one, therefore we should think it to be true (Lewis, 1986a: 3-4).¹⁶

Despite criticisms, Lewis (1986a: 17) supported the view that his version of modal realism in possible worlds enabled him to do metaphysics, instead of pure semantics or mathematics. Notably, Lewis (1986a: 4) recognized certain dangers in resorting to a set theoretical interpretation for possible worlds talk in modal realism: "May be the price is higher than it seems because set theory has unacceptable hidden implications – may be the next round of set-theoretical paradoxes will soon be upon us". This Lewisian argument (1973b: 89) comes as a critique of Quine's "generalization" for settling the possible worlds notion in reductive terms, which, for Quine, means "in correspondence with certain mathematical structures representing the distribution of matter in space and time".¹⁷ The Lewisian thesis that possible worlds do not differ in kind from our world is consistent with the Lewisian explanation that mathematical entities do not constitute possible worlds. As Lewis (1986a: 4) explains, his theory is only a "parallel with mathematics", which serves his philosophical purpose. Lewis (1973b: 90) argues:

"I cannot believe (though I do not know why not) that our own world is a purely mathematical entity. Since I do not believe that other worlds are different in kind from ours, I do not believe that they are either. What is interesting is not the reduction of worlds to mathematical entities, but rather the claim that the possible worlds stand in a certain one-to-one correspondence with certain mathematical entities. Call these ersatz possible worlds."¹⁸

¹⁶ Lewis (1986a: 5) also states that any doubts against the "controversial ontology" of his modal realism are not the subject of his work *On The Plurality of Worlds*.

¹⁷ For instance, taking worlds in Quine's view as worlds "where space-time is Euclidean and four-dimensional, and where there is only one kind of matter and no fields"; or other worlds, where "there might be scalar, vector, or tensor fields independent of the distribution of matter", or where "there might be more than one kind of matter, or more or less density of matter" (Lewis, 1973b: 89). Lewis (1973b: 89-90) argues instead against such a generalized construction of possible worlds, even when it may account for overlooked possibilities, such as: non-Euclidean space-time; scalar, vector, or tensor fields independent of matter's distribution; more than one kind of matter, or more or less density of matter. As Lewis (1973b: 90) explains: "I do not, of course, claim that these complicated mathematical entities *are* the possible worlds".

¹⁸ However, Lewis (1986a: 95) accepts that universal numbers and mathematical objects are parts of this world, hence they can be called 'actual' in his modal realist terms. Sider (2009: 18) has referred to Lewisian realism in modal metaphysics as "ontological realism", in scope of his use of the term to define the metaphysical claim that quantificational structure is included in the world's distinguished structure (Sider, 2009: 26).

Several arguments have been put forth, also by Lewis himself (1986a: viii), that "modal realism" is not the most appropriate terminology for defining the Lewisian system of thought in terms of the notion of possible worlds (Nolan, 2005: 54).¹⁹ Besides "extreme" or "genuine" (Divers, 2002; Lewis, 1983c, 1986a; Lycan, 1991; Miller, 1993), Lewisian modal realism has also been characterized as: i) "modal reductionism", because of its seemingly reductivist treatment of propositions, properties, and possible worlds on a set theoretic basis (Plantinga, 2003); ii) "modal fictionalism" (Nolan, 1997), because it is modal realist about possible worlds and not about the modals of possibility and necessity (Nolan, 2005: 54). I return to these arguments in the last evaluative **section 8** of this thesis. Nevertheless, from now on I use the established term of modal realism for Lewis's philosophico-logical project, whilst bearing in mind the differences between Lewisian modal realism and other positions from the modal realist strand that Lewis calls 'ersatzism'.²⁰

¹⁹ Lewis acknowledges that, with his modal realism, he is doing metaphilosophy. However, he also carefully argues that his modal realism produces metalogical results, which are applicable in a set theoretical context (Lewis, 1986a: 20), although possible worlds are necessary for the substantive theory, but not the metalogic. As Lewis (1986a: 2) argues:

"When I say that possible worlds help with the analysis of modality, I do not mean that they help with the metalogical 'semantical analysis of modal logic'. Recent interest in possible worlds began there, to be sure. But wrongly. For that job, we need no possible worlds. We need sets of entities which, for heuristic guidance, 'may be regarded as' possible worlds, but which in truth may be anything you please. We are doing mathematics, not metaphysics. Where we need possible worlds, rather, is in applying the results of these metalogical investigations. Metalogical results, by themselves, answer no questions about the logic of modality. They give us conditional answers only: if modal operators can be correctly analyzed in so-and-so way, then they obey so-and-so system of modal logic. We must consider whether they may indeed be so analyzed; and then we are doing metaphysics, not mathematics".

²⁰ Lewis uses this term, instead of the more neutral term 'abstractionism', which has been employed by van Inwagen (1986). Lewis himself (1986a: vii) defines "the programme of ersatz modal realism" as the one "in which other worlds are to be replaced by 'abstract' representations thereof". Drawing from his counterpart theory, Lewis (1983c: xi) distinguishes his version of "extreme modal realism" from other modal realist approaches by offering the following description: "Extreme modal realism, according to which there are many unactualized possible individuals, and according to which the actual individuals do not differ in kind from the unactualized ones". See **Appendix I**, "Objections and Replies", 3, where I attend to Lewis's response regarding modal reductionism and modal fictionalism, in respect to different modal realist strands.

3

Lewis's philosophical method of analysis

Taking conceptual analysis as an established general method of philosophical analysis in analytic philosophy, I here present Lewis's own general philosophical method of analysis, while also explaining its application in an example from Lewis's value theory, namely his dispositional value theory.

Conceptual analysis in analytic philosophy consists partly in analyzing theoretical terms and partly in producing claims about subject matter, either seemingly obvious, or difficult to dismiss by specialists in complementary areas, such as in linguistic analysis.²¹ For Nolan (2005: 222-223), in borrowing from linguists, analytic philosophers tend to follow this analytic method as informal hypothesis, rather than for testing, through empirical and other means, a hypothesis in order to survey linguistic behavior. For instance, likewise in his preliminary exposition of his possible worlds thesis in *Counterfactuals* (1973), Lewis (1973b: 91) argues that the construction of ersatz worlds, or possible worlds in his 'genuine' modal realist view, may be influenced by a reconsideration of modal opinions concerning current physics, rather than by verification on the basis of empirical scientific results.²²

In his method of philosophical analysis, Lewis (1970: 427) argues for a "vindication of theoretical terms" following the Ramsey-Carnap tradition. This

²¹ This method may follow an established compositional process of organization in analytic philosophy, which describes the transition from natural language, through a sort of translation, to a logical language, and, through a sort of interpretation, to logical models; see (Gamut, 1991: 149). See also (Bennett, 1964) and (Sellars, 1980).

²² For Lewis (1986a: 59), the theoretical model of possible worlds is not doctrine, but useful terminology, which combines "set theory and modal realism in structured and unstructured versions of properties, relations and propositions". Lewis (1986a: 234) is aware that new terminology does not make up for a new theory, but rather works on a modification of a previous one, in this way following the Ramsey-Carnap tradition, as explained in this section.

tradition suggests not doing away with theoretical entities for the sake of newly introduced terms, but defining them by following a general method; for example, in the philosophy of science, we understand them as theoretical terms, introduced by a given theory at a given stage of the history of science (Lewis, 1970: 427-428).²³ As I show in the following **section 4**, Lewis offers examples when advocating the use of theoretical terms in his value theory and in his metaphysics, where he adopts respectively: i) Carnap's method of using quantitative concepts for describing moral value functions to show, like Carnap (2017 [1958]: 189), that it is not feasible to do so; ii) a restatement of Ramsey's 1928 theory of lawhood to explain, in his Lewisian metalinguistic theory of conditionals, why and in which conditions lawhood is a contingent property (Lewis, 1973: 73-75).²⁴ Furthermore, Lewis (1970: 428) defines theoretical entities as hypothetical, which do not need to be abstract or be understood strictly by means of theoretical terms.²⁵ Yet in *Counterfactuals*, Lewis (1973: 70-72) contests the wide applicability of such entities: for instance, in his metalinguistic theory of conditionals regarding factual premises; while he also implies that possible worlds might be such entities (Lewis, 1973: 85).

Arguably, Lewis treats philosophical analysis in a general manner, which is consistent with a non-strictly linguistic approach in analytic philosophy. As Lewis (1986: 4) argues in his metaphysical thesis of possible worlds: "We make languages and concepts and descriptions and imaginary representations that apply to worlds".²⁶ This methodological approach is generalized when applied to the

²³ The term 'theory' is here employed generally, as the body of a theory to include a set of its statements, which could be referred to as laws, propositions, regularities, theorems, or rules.

²⁴ Ramsey's 1928 theory of lawhood is stated in Lewis (1973b: 73): "[...] laws are 'consequences of those propositions which we should take as axioms if we knew everything and organized it as simply as possible in a deductive system'". Ramsey later revised this original hypothesis; see (Lewis, 1973b: n73). Lewis (1973b: 73) restated Ramsey's hypothesis as follows: "A contingent generalization is a law of nature if and only if it appears as a theorem (or axiom) in each of the true deductive systems that achieves a best combination of simplicity and strength". From this revision follows that "a generalization is a law at world *i*, likewise, if and only if it appears as a theorem in each of the best deductive systems true at *i*" (Lewis, 1973b: 73).

²⁵ Lewis (1970: 427) argues that we believe in a hypothetical entity "only because its existence, occurrence etc. is posited by some theory – especially some recent, esoteric, not-yet-well-established scientific theory". Nolan (2005: 213) refers to such type of analysis as the "Ramsey-Carnap-Lewis" analysis of theoretical terms, which he describes as follows: "When we have a theory of some phenomenon, we can put together all the claims of the theory, and turn them into a generalisation that serves the same purpose as the original theory, but contains none of the problematic expressions to be analyzed".

²⁶ Several interpretations consider Lewis and Stalnaker to be philosophers dealing with a conventionalist treatment of conversational linguistic variables; as argued by Thomason (2011: 830): "[...] as information that is dynamically updated and maintained by the participants in the conversation". Their treatment of conversational linguistics has been significant for non-monotonic linguistics. Furthermore, Lewisian conditional logics have helped the development of contemporary theories of generics in linguistics and non-monotonic reasoning (Van Rooij and Schulz, 2011: 849),

problem of analyticity, which is a problem of the philosophy of language concerning the truth conditions assigned to a sentence (Lewis, 1969: 175). In his essay "General Semantics" (1972), Lewis (1983d [1972b]: 189) characteristically contrasts a general approach to semantics with an empirical one:

"It is not my plan to make any strong empirical claim about language. To the contrary: I want to propose a convenient format for semantics general enough to work for a great variety of logically possible languages. This paper therefore belongs not to empirical linguistic theory but to the philosophy thereof."

In differentiating between general referential descriptions of languages and grammars as abstract semantic systems of symbols, which are "associated with aspects of the world" (Lewis, 1983d [1972b]: 189), and psychological or sociological descriptions particular to persons or populations, Lewis insists on the semantic method of specifying truth conditions. As Lewis (1969: 1-2) explains in *Convention*:

"To say this is not to say much. It is not to portray language in the image of a calculus, precise and rigid. It is not to uphold "correct" speech against colloquial, or vice versa. It is not to say that all the languages we can think of are equally good, or that every feature of a serviceable language might just as well have been different. It is not to say that necessary truths are created by convention: only that necessary truths, like geological truths, are conventionally stated in these words rather than in those. It is not to exalt the powers of convention as some "conventionalist" philosophers do, but only to insist that it is there. The platitude that there are conventions of language is no dogma of any school of philosophy, but commands the immediate ascent of any thoughtful person – unless he is a philosopher."

Despite complying with the general approach to philosophical analysis, Lewis's metaphysical pursuit in his *Counterfactuals* and *The Plurality of Worlds* has not prevented him from working on the metaphilosophical level: on "producing a theory of producing a theory" of something, instead of producing a theory of something, as Nolan (2005: 213) argues. Such methodological approach requires a method of reasoning, which starts from certain non-contingent specified premises (Lewis, 1986a: 115). Using a combination of informal reasoning and reasoning from mathematical axioms, Lewisian modal realism (1986a: 115) is exemplary of such twofold philosophical analysis, in its distinction between the modal case, which proceeds from highly informal reasoning and imaginative experiments, and the

which is "higher order and complex reasoning", where "fool-proof strategies are scarce" (Van Benthem, 1996: 295).

mathematical case, which proceeds from more or less rigorous, and sometimes informal, reasoning from mathematical axioms, or from already accepted general principles (Lewis, 1986a: 113).

I show in **section 5** that Lewis's folk theory of so-called 'moral science' is an interesting example of Lewisian modal realist interpretation, which offers more or less rigorous applications of modal reasoning in Lewis's conventionalist and modal metaphysical approaches. By appealing to folk psychology,²⁷ Lewis (2000 [1996b]: 63) argues that we modalize in everyday life following an informal method of modal reasoning; for example, when we are pleased that an undesirable outcome has not turned out to be actual. Lewis (1986a: 7) further argues that such reasoning is applied to his possible worlds analysis, in his use of the notion of impossibility in logical terms to make a counter comparison with the notion of contingency.

In consistence with his treatment of the notion of analyticity, Lewis's general approach to semantics is also recognizable in his value theory. One such application is Lewis's early schema of a general theory of persons (1974), seeking a twofold interpretation of personhood, based on a person's language and system of beliefs and desires.²⁸ Lewis's later defense (2005) of his moral theory of dispositional value against the strand of fictionalism and strands of projectivism in ethics provides another such application. Departing from the canonical position in ethics of recognizing genuine values in terms of necessity *simpliciter*, or metaphysical necessity, Lewis's method of analysis in his value theory complies with his view of philosophical study as analytic, even if not obviously so, as I go on to explain in the following subsection.

²⁷ I here use the established term of folk psychology, which has been given various definitions. Notably, Allen and Bekoff (1997: 65-66), who propagate that folk psychology constitutes a prototheory of social behaviour, argue that in current debates "[...] it consists of a rather loose set of generalizations about mind and behaviour that are reflected in the things that normal adult humans say about mental states and action".

²⁸ Lewis (1974a: 334) puts forward the standard problem in analytic philosophy of radical interpretation, as addressed by Quine and Davidson, to define the constraining principles by which it can be solved, as "the fundamental principles of our general theory of persons". Lewis (1974a: 336-339) lists these principles as: the principle of charity, the rationalization principle, the principle of truthfulness, the principle of generativity, the manifestation principle and the triangle principle. Lewis (1974a: 340-341) argues that his method borrows from Davidson's analysis of radical interpretation, while acknowledging that Davidson's problem is not the same as his problem, but it can rather be treated as a sub-problem of his. For Davidson's position on radical interpretation and radical translation, see (Davidson, 1973, 1984).

3.1 The example of Lewis's dispositional theory of value

Lewis exemplifies his general analytic method in his dispositional theory of value. Here, Lewis (2000 [1989a]: 85-86) defines his value theory analysis by means of the modal status of the equivalence between value and what one is disposed to value. In virtue of this definition, he describes his methodological approach as both "unobviously" and "equivocally" analytic;²⁹ as Lewis (2000 [1989a]: 85) argues, such equivalence "is not obviously analytic; it is not even obviously true". Lewis (2000 [1989a]: 85-86) illustrates the analytic method of his dispositional value theory, which poses the problem of 'unobvious' analyticity, in his analysis of the following propositions: whether all A's are indeed A's, which is a tautology, and whether all A's are B's, which is not tautological. Lewis argues that these 'unobviously' analytic statements are in fact intelligible analytic statements. Additionally, Lewis (2000 [1989a]: 86) shows that the 'equivocal' analyticity of his dispositional value theory takes semantic variation into consideration without turning away from analyticity.

For instance, it is 'equivocally' analytic the question concerning the "disambiguation", or "precisification" – meaning dismissing ambiguity, or making precise – of the analytic problem of what is the height constitutive of a 'mountain'. Nonetheless, the question opens a conversation to the answer that may not be analytic: for example, the concept of a mountain can stand for a hill, or for Mount Everest.³⁰ Whether there is semantic variation and indecision,³¹ or simply lack of truth value, in Lewis's dispositional theory of value (2000 [1989a]: 88), which is seminal for his folk theory, one can come up with counterexamples. As Lewis (2000 [1989a]: 86) explains, "something may be analytic under one disambiguation, but not another", "under one precisification, but not another". Taking examples of analytic ambiguity, Lewis (2000 [1989a]: 87) argues that, in hypothesizing that his dispositional theory of value is analytic, both 'unobviously' and 'equivocally', provides an approximation of the canonical method of analysis in Lewis's value

²⁹ Nolan (2005: 183) explains that Lewis's dispositional value theory is meant to be an analytic truth in an 'unobvious' way, which implies that one could use the word "value" without agreeing with Lewis's analysis; while being 'equivocally' analytic means that it remains ambiguous or undetermined what people mean by the expression "value".

³⁰ See also Putnam's argument for divergent interpretations of such analytic claims (1983: 118-120) also by referring to Lewis's physicalist position.

³¹ Lewis (1983e [1975a]: 188) accepts what he calls "unsharp analyticity", in cases where it cannot be decided whether a sentence is true in a possible, but bizarre, world. To clarify his notion of 'unsharp analyticity' without a possible worlds analysis, Lewis (1983e [1975a]: 188) admits to conventions of truth and trust in a cluster of similar languages, rather than in a single language.

theory, whilst explaining one's belief, also in folk terms, that values equate with what one is disposed to value.

Notably, in his general semantics analysis, Lewis (1972: 170) distinguishes between two topics, which, in his view, should not be conflated with one another:

"[...] first, the description of possible languages or grammars as abstract semantic systems whereby symbols are associated with aspects of the world; and second, the description of the psychological and sociological facts whereby a particular one of these abstract semantic systems is the one used by a person or population."

Together with his dispositional value theory, Lewis's theory of general semantics is useful for understanding his interpretation of folk ethics in its differentiation from other interpretations, such as those considering psychological and sociological facts. As I argue in **section 4**, in Lewis's folk ethics, value theory only approximates methodologically canonical ethics, which explains why folk would take values to be what the folk is disposed to value. As I show in view of Lewis's folk ethics, or "moral science" (2000 [1996b]: 59), when general semantics is applied to Lewis's dispositional value theory including his population semantics, there is an evident difference between a general semantics analysis and say a psychological and social analysis. This distinction primarily enables Lewis to offer a value theory that is both 'equivocally' and 'unobviously' analytic, hence Lewis can follow the general approach to semantics by means of the treatment of truth conditions; as he does, for instance, in his value theory example of the disambiguation or precisification of the word 'mountain', where there is relativity in the application of analyticity.

4

Value theory: dispositional theory of value as a solution to the problem of the anti-Humean desire as belief thesis

In this section, I consider Lewis's groundwork for his ethics and social philosophy in my analysis of the Lewisian value theory scheme. I show that, against the anti-Humean desire as belief thesis, Lewis proposes his dispositional theory of value. Drawing from Lewis's seminal value theory papers "**Desire as Belief**" (1988), "**Dispositional Theories of Value**" (1989), and "**Desire as Belief II**" (1996), I analyze Lewis's model-theoretic interpretations of the anti-Humean Desire-As-Belief thesis, which is determined probabilistically and conditionally, in view of Lewis's dispositional value theory as equivocally and unobviously analytic, but also as conditionally relativist.

The Lewisian dispositional value theory scheme is an amendment of the Humean thesis that an agent's motivation derives from his or her disposition to act upon desire according to belief. In view of this scheme, Lewis (2000 [1989a]: 68) defines value as what one is disposed to value. Hume (2007: 35-36) associated beliefs with sentiments, thus avoiding a presumed idealism of belief and rather supporting a conceptual modal and cognitive interpretation of belief, which is contrary to a fictionalist one. Regarding the influence upon an agent's motivation for belief and desire, which, for Lewis (2000 [1988]: 49), has a relevant, yet not strictly constraining, impact on decision theory, the anti-Humean thesis prioritizes belief, while it either conflates belief with desire, or necessarily aligns beliefs with corresponding desires (Lewis, 2000 [1988]: 43). In his interpretative analysis of the anti-Humean thesis, Lewis shows that there is a collision of the anti-Humean Desire-as-Belief thesis and standard decision theory (Lewis, 2000 [1988]: 45).

On the basis that there is a lack of informative evaluative analysis of the content of belief on the anti-Humean strand, to explore the anti-Humean thesis, Lewis (1988) uses probability conditionals in his analysis of belief content. Lewis (2000 [1988]: 48) applies the 'invariance assumption', which supports that there is no change in the expected value of a given proposition "in response to an exogenous redistribution of credence over a partition $E_1, \dots, E_n$ ", if the given proposition is compatible with only one of the E_i 's, in order to formalize desire as belief propositions equating values to unconditional or conditional beliefs. Lewis (1988, 1996b) also applies model-theoretic interpretations for the anti-Humean theses Desire-As-Belief (DAB) and Desire-As-Conditional-Belief (DACB), with the second thesis refining the first one (Lewis, 1996b), which concern respectively the unconditionality or conditionality of belief. DAB, which is the thesis that an agent desires something when the agent believes it would be good, is formulated probabilistically, using a credence function C and an expected value function V , which are formulated by Lewis (2000 [1988]: 46) as follows: taking any proposition A and any partition E , with partitions $E_1, \dots, E_n$ being a set of mutually exclusive and jointly exhaustive propositions, for any proposition A and any partition $E_1, \dots, E_n$,

$$(1) \quad C(A) = \sum_i C(AE_i) = \sum_i C(A/E_i)C(E_i)$$

$$(2) \quad V(A) = \sum_i V(AE_i)C(E_i/A) = \sum_i \frac{V(AE_i)C(A/E_i)C(E_i)}{C(A)}^{32}$$

The above formulas comply with the definition of DAB as: "the value of a proposition that might come true in several alternative ways is an average of the values of those several alternatives, weighted by their conditional credences" (Lewis, 2000 [1988]: 46-47). With the introduction of Lewis's notion of 'probability kinematics' (2000 [1988]: 47), which is a term standing for the change in someone's state of credence and value from an initial state, Lewis considers the impact of the change of credences on resulting changes of valuations to expected values. In the case of imposing DAB as a new constraint on decision theory, Lewis (2000 [1988]: 49) argues by applying DAB to an example of a given agent, that the agent can simultaneously change his belief and his desire due to a change by so-called 'probability kinematics', in this way demonstrating that credence and expected

³² In his later essay "Desire as Belief II", Lewis (2000 [1996b]: 56) applies the following revised definitions: $C(A) = \sum_i C(AE_i)$ and $V(A) = \sum_i V(A_i)C(A_i/A)$.

value of any proposition A can change simultaneously. Likewise, considering degrees of value, in the example of an agent taking a moral value, such as goodness, to admit of degree (Lewis, 2000 [1988]: 52), the agent can simultaneously change his belief and the component of desire deriving from this belief regarding the degree of goodness; hence, there is a correlation between belief about the degree of a value and the corresponding component of desire drawn from the belief.

To reach the above evaluative results, Lewis (2000 [1988]: 48) revises the 'invariance assumption' that if a proposition A is compatible with only one of the elements of a partition E, the value of A does not change. This revision is formulated as follows. Taking x as a variable, which can be attributed some range of values, and d numbers measuring the distribution of change, we have:

$$(5) \quad V_x(A) = V(A) \frac{[1+rx]}{[1+px]}, \text{ where } r = \frac{\sum_i V(AE_i)C(A/E_i)d_i}{V(A)C(A)}$$

For this reason, Lewis (2000 [1988]: 50) revisits the thesis of the 'probability conditional', that probabilities of conditionals are conditional probabilities, or equal the corresponding conditional probabilities. Taking P as the probability function, A any proposition and C the consequent, we have:³³

$$P(A \rightarrow C) = P(C/A)$$

Applied to Lewis's analysis of DAB (2000 [1988]: 50), the above provides the following formulas. Taking C as the credence function, A any proposition, G an agent's goal, we have:

$$(I4) \quad C(A \rightarrow G) = C(G/A)$$

$$(I5) \quad C_x(A \rightarrow G) = C_x(G/A)$$

Lewis proves that we cannot give a sense to the entailment ($\rightarrow$), which does not work, for all C, A and G, for (I4), neither, for all A and G, for (I5), even if we fix C, an originating partition and distribution.³⁴ Equally, Lewis (2000 [1988]: 50) demonstrates that the probability conditional does not hold when degrees of value are introduced. In view of the results against the probability conditional, Lewis's argument (2000 [1988]: 51) concludes on its unsuccessful application on the case of a single minded agent, to further argue, against the anti-Humean DAB, for the preservation of the 'invariance principle' and for subcases to be taken as maximally

³³ See also Lewis's (1976a) examination of this thesis supported by Stalnaker.

³⁴ See also a similar direct proof that (I4) does not hold for fixed C, fixed partition and distribution, and all x, A and G, which Lewis (2000 [1988]: n51) presents by examining only a single conditional, arguing that "generality over consequents is not much used".

specific in all relevant propositions regarding degrees of value (Lewis, 2000 [1988]: 53-54).

In his previous paper, "Probabilities of Conditionals and Conditional Probabilities" (1976), examining the same problem of the probability conditional drawing upon Stalnaker's conditional introduced in *Counterfactuals*, Lewis (1986d [1976]: 149) suggests that entailment ($\rightarrow$) itself is taken as a probability revision conditional for a given method, if and only if the following formula holds in general. Taking PA as the revision obtained by this method, we have:

$$(20) \quad P(A \rightarrow C) = PA(C), \text{ if } A \text{ is possible.}$$

Here Lewis (1986d [1976]: 150) argues for methods of revision for both worlds and probability functions. Lewis (1986d [1976]: 151) offers an example of such a method, taking an opinionated probability function, by means of a realist approach, as "the revision of certain special probability functions that stand in one-to-one correspondence with the worlds". Such probability function represents an agent's belief of absolute certainty about the actuality only of a world W , where probability function P concentrates all the probability. Therefore, W is the belief world of P , that is $W(A)$ equates with $P(A)$. $W(A)$ is formulated as a set, where each of the values 1 and 0 are ascribed respectively to the truth or falseness of a given proposition A . For Lewis (1986a: 151), this method preserves opinionation, which allows a given agent to hold a firm opinion: revisiting formula (20), with P being the opinionated probability function, entailment ($\rightarrow$) is not a probability revision conditional and $PA(C)$ does not equate $P(A \rightarrow C)$. In this case, Lewis (1986a [1976a]: 151-152) offers a method of revising worlds, by equating $PA(C)$ with $WA(C)$, where WA is the revision obtained and entailment ($\rightarrow$) is the Stalnaker conditional for the method of revising worlds.

In the above described formalization and analysis of the content of desire as belief with respect to probability conditionals, Lewis (2000 [1996b]: 58) examines the "anti-Humean moral science", which, amongst others, suggests discarding the 'invariance principle', but also may rely on anti-Humean metaphysics of modality. Characteristically, in his critique of anti-Humeanism, Lewis (2000 [1996b]: 61) takes one of its possible versions, that, despite their contingency, desires are necessarily aligned with suitable beliefs. By applying different interpretative approaches to the anti-Humean theses, including his model-theoretic analysis of DAB and its revision by DACB, Lewis (2000 [1996b]: 67) concludes on the anti-Humean counterargument

that these approaches: either collapse into contradiction or triviality; or they are the wrong form of anti-Humeanism, as in DACB, where contingent desires are not necessarily aligned with beliefs (Lewis, 2000 [1996b]: 64); or they are not at all anti-Humean, as in the DAB equation, which, however, does not hold in virtue of any interesting connection between a desired and a believed proposition, but only in virtue of a proposition A° about the proposition A holding the right credence, which equals $V(A)$, thus, there is nothing to suggest that the proposition A° about proposition A is objectively good (Lewis, 2000 [1996b]: 66).

While upholding Humeanism to define value as "the desire to desire", which does not equate 'valuing' with 'desiring', Lewis's solution to how value is determined (Lewis, 2000 [1989a]: 70-71) follows an *approximation* to the canonical method, which defines value by establishing that: "something of the appropriate category is a value if and only if we would be disposed, under ideal conditions, to value it" (Lewis, 2000 [1989a]: 68). Lewis's schematic dispositional value theory shows that, under the canonical interpretation, values are determined conditionally and by means of informal hypothesis, or an "imaginative exercise" (Lewis, 2000 [1989a]: 78). As Lewis (2000 [1989a]: 83-84) explains, values are determined in conditions, which are ideal, or "extravagantly ideal", for these values, through namely 'imaginative acquaintance' taken as informal hypothesis. In applying informal hypothesis as a method of establishing how values are determined, we easily come to the realization that, even amongst a population of psychological clones, who are all disposed to value alike, there will be diversity in valuing, because people would have been led to different imaginative exercises in determining their values. More importantly, for Lewis (2000 [1989a]: 84), the canonical method of establishing one's values, whatever the values to be determined are, absolute or relative, also determines whether someone ascribes oneself to the appropriate 'we'.³⁵ Lewis (2000 [1989a]: 81) rather suggests that his dispositional value theory approximates the canon by locating valuing in worlds – say a world like *ours*, with reference to the indexical 'we' – and by accepting the contingency of value as an aspect of his theory due to the contingency of psychology (Lewis, 2000 [1989a]: 88-89).

Moreover, in offering an interpretation to his solution, Lewis draws upon his method of philosophical analysis in value theory, as unobviously and

³⁵ However, for Lewis (2000 [1989a]: 84), the appropriate 'we' would not apply to 'elitist' versions, where 'we' means the 'best-qualified of us' or maybe 'the most normal of us'.

equivocally analytic, and the application of modal reasoning.

1. To begin with, Lewis (2000 [1989a]: 68) broadly describes his account of the dispositional value theory by virtue of "a naturalistic metaphysics", because the theory advances an analytic definition of value, but also because it is reductive.³⁶ As Lewis (2000 [1989a]: 68) mentions:

"It invokes such entities and distinctions as we need to believe in anyway, and needs nothing extra before it can deliver the values. It reduces facts about value to facts about our psychology."

Lewis (2005: 230) further explains that his dispositional value theory "is a kind of analytic naturalism, though I admit that (like most interesting analyticity) it may be analytic under some and not all legitimate resolutions of semantic indeterminacy";³⁷ hence, in Lewis's view (2000 [1989a]: 92), it is equivocally analytic. Lewis (2000 [1989a]: 68) also points at the subjectivity of his value theory in the broad sense, since it analyzes value with regard to our attitudes, yet without being subjective in the narrow sense, which implies that whatever one may think on the subject of value is automatically true, or that there is no truth at all about values. Neither does his value theory imply a strict dispositionalism, meaning that if we had been disposed differently, values would have been different themselves. The latter description points at the unobvious analyticity of Lewis's value theory, in view of which Lewis problematizes the question of the realism of values, without quickly settling the question in one way or another. For instance, in choosing between two versions of moral language to describe value, the language of absolutism and the language of relativity, Lewis (2000 [1989a]: 85) commits to conditional relativity, "if need be, but absolute otherwise".

In his interpretation of the approximation to the canonical method, Lewis takes on an analysis of the indexical 'we'. Lewis (2000 [1989a]: 82-83) accepts that, on the one hand, "an absolute version of the 'we' refers to all mankind", in which case a value is a value *simpliciter*, or universal value, meaning that "everyone, always and everywhere, is disposed under ideal conditions to value it". On the other hand, he points at the recognition that all mankind is alike besides the diversity of valuing due to cultural differences, which yet does not provide counterevidence to the recognition of universalism to a certain extent. Nevertheless, Lewis (2000 [1989a]:

³⁶ Another interpretation can here be offered by Searle's account (2009) of naturalizing language.

³⁷ See also **Appendix I**, "Objections and Replies", 2.

83) also recognizes that "the language of absolutism does not prevail the language of relativity": 'we' also consists of an indexical, which refers to a population consisting of the valuer or agent and those like him.

2. Lewis introduces modal reasoning in his value theory to advance his analysis of the ontological status of values. As long one accepts one version, what Lewis (2000 [1989a]: 91) calls the original version of his dispositional value theory, for which value is contingent, Lewis (2000 [1989a]: 90) argues that our *dispositions* to value are 'necessary' and not contingent, but that *values themselves* are not 'necessary'. For if the latter was the case, either, strictly speaking, there are no values, or we fall into the error theory, or an amended version of the dispositional value theory, requiring that values are all that we want them to be. Still, for Lewis (2000 [1989a]: 90), we are not *necessarily* disposed to value certain values *per se*, neither in nomological, nor in deontic terms (Lewis, 2000 [1989a]: 89), while it is also implied that there are no values in terms of necessity *simpliciter*. In Lewis's view (2000 [1989a]: 92), this becomes evident specifically upon our amendment of the dispositional value theory, which consists in his refined second version, by introducing the modal 'necessarily' to discuss whether there are values that we are 'necessarily' disposed to value and how we ascribe the term 'value'. What is Lewis's suggestion? Turning to folk psychological statements (Lewis, 2000 [1989a]: 90-91), which have a certain conditional necessity, Lewis (2000 [1989a]: 93) suggests that here we can get an indication of mental states, which deserve folk psychological names and could serve as approximations of what we value in ideal conditions, but also as definitive of the plurality of values, which we are disposed *de facto*, or as a standard, to value.

Lewis (2000 [1989a]: 80) continues with applying modal reasoning to exemplify his notion of the disposition to value as conditionally relative by an informal hypothesis, in which he uses the figure of the ideal responder to argue that, in the case of he or she being an ideal valuer, we need to replace the ideal valuer with an ideal balancer of values. Lewis (2000 [1989a]: 80) argues:

"Genuine values might be unattainable, or unattainable without undue sacrifice of other values. An ideal balancer of values needs thorough knowledge of the terms of trade. An ideal valuer may be better off without it. Our present business is not with the balancing, but with the prior questions of *what values there are to balance*."

Despite posing the problem of the ontology of values, Lewis does not say what values are, but rather settles the question in reference to and in the context of a possible agent as a valuer and ideal responder, whom either has or lacks putative values, or as ideal balancer of values, and a possible world. As I argue in **section 6**, when Lewisian modal realism is applied in deontic terms, for instance in Lewis's sphere semantics and counterfactual conditionals, it does not abide to an absolute external standpoint, or a theological viewpoint.³⁸ Likewise, in his dispositional theory of value, Lewis equally problematizes the notion of a possible valuer or agent, who does not refer to any absolute standpoint. Lewis (2000 [1989a]: 93) places his theory of value between irrealism and realism, showing that values do not strictly, but only loosely, exist.

³⁸ Commenting on the many-worlds hypothesis borrowed from quantum mechanics, Putnam (1990: 8-11) argues that there has been an unsuccessful attempt to introduce the omniscient observer as the God's eye view, to counteract the quantum physics' splitting effect of branching. For Putnam, this scientific effort had an influence on the philosophy of logic and on metaphysics.

5

Lewis's folk 'moral science'

I now analyze Lewis's folk 'moral science' in relation to Lewis's value theory scheme, specifically his dispositional theory of value. I undertake this analysis in view of Lewis's efforts to systematize folk theory by his philosophico-logical project of modal realism. Lewis proposes to take language as semantics integrated into language usage in a population and in social interaction, as I mentioned in **subsection 1.2**. On this basis, I discuss: Lewis's conventionalist approach of semantics in a population in **subsection 5.1**; and Lewis's modal metaphysical approach of semantics in a population according to worlds in **subsection 5.2**.

As has been discussed in **section 4**, the Lewisian commitment to dispositional value adheres to the conditional relativity of values established by a method, which only approximates the canonical in its dismissal of the canon's inherent idealism. I also showed that, in his analysis of the anti-Humean accounts of Desire-As-Belief and of dispositional value, Lewis defends the Humean against the anti-Humean thesis. I now concentrate on Lewis's characteristic analysis of "folk moral science", which Lewis (2000 [1996b]: 59) defines as a combination of folk ethics and folk psychology, to discuss whether it can be explained in view of the aforementioned Lewisian commitments. I argue that, by turning to folk moral psychology, Lewis builds upon his Humean dispositional value theory, but also his view of the conditional relativity of value according to populations. Besides other concerns, Lewis's folk 'moral science', but also his thesis on the plurality of worlds, which I discuss in the following **section 6**, draw upon Lewis's value theory scheme, specifically his dispositional theory of value, as outlined in the previous **section 4**. As I have argued, Lewis's value theory offers an example of a theory that is unobviously and equivocally analytic.

The Lewisian account of folk moral science is central to Lewis's analysis of moral philosophical questions, which Lewis (1969: 195) situates relatively to the problem of the use of language as semantics in a population, as I discussed in section 3. Lewis (2000 [1996b]: 59) appoints part-ethical and part-psychological meanings to folk theoretical vocabulary, as well as an operative role in its introduction of theoretical terms, which is close to, but not the same as or equal to, the respective role of scientific theoretical terms; hence, his reference to folk moral theory as "folk moral science" (Lewis, 2000[1996b]: 59). As Lewis (2000 [1996b]: 58) argues:

"[...] tacitly known folk theories may introduce terms in much the same way that scientific theories do; and, in particular, that our ordinary mental vocabulary consists of the theoretical terms of commonsensical 'folk psychology'. Belief and desire, among others, are the states that occupy certain folk-psychological roles. And again, when it comes to occupying a role and thereby deserving a name, near enough is good enough. Folk psychology needn't be flawless!"

Lewis (2000 [1996b]: 60) credits folk moral psychology with a certain degree of objectivity for describing a value scheme, in its approximation to a scientific, or canonical, system. For this reason, when Lewis mentions folk theoretical postulates (2000 [1996b]: 59), which say either what is universally desired, or how desires are aligned with beliefs, he uses both psychological and ethical vocabulary to define folk moral theoretical terms by approximation to scientific terms (Lewis, 2000 [1989b]: 92). Interestingly, Lewis (2000 [1996b]: 60) believes that, despite the schisms in folk ethics, or probabilistic interpretations of consensus, or other obvious relativistic concerns, one should care about folk moral psychology in its signification of approximations of objective value and ethical reality (Lewis, 2000 [1996b]: 58-59). Lewis (2000 [1996b]: 60) explains that, if one does not care about ethical reality and objective value in folk theoretical terms, one's inner state will fail to deserve folk theoretical names. This Lewisian suggestion does not mean, or even imply, that Lewis appoints folk ethics with any philosophical absoluteness.

In his analysis of the anti-Humean thesis, Lewis (2000 [1989a]: n.69) brings forth another example from canonical ethics by mentioning the moral philosophical term '*besire*', which describes "an attitude of valuing":³⁹ "a special

³⁹ For the notion of motivation, Lewis draws from Anscombe's essay on intention. Anscombe (1981: 77) recognizes a philosophical distinction between motives and intentions in action as referring to

kind of attitude that is both a belief and a desire and that motivates us, without benefit of other desires, in just the way that ordinary desires do". Yet when Lewis (2000 [1989a]: n70) attempts to integrate the notion of 'desire' in his interpretations of folk psychology, he encounters difficulties because of his view that a belief system is likely to change after the impact of new information. In this view, he offers an explanation in his model-theoretic analysis of Desire-As-Belief about why beliefs may evolve in one way, and desires, if instrumental, in another way, while experience may equally influence our evaluation and definitions of the intrinsically good. Lewis here shows, as I pointed out in **section 4**, that when the anti-Humean thesis is applied on folk moral science, the Humean thesis is, in fact, reasserted.

We also saw in **section 4** that, in analyzing the notions of belief and desire in modal realist terms, to avoid the difficulties posed by anti-Humeanism, Lewis proposes his dispositional value theory, which he describes, apart from naturalistically metaphysical and subjective, by using specific theoretical terms. Lewis's terms for his value theory (2000 [1989a]: 69), which are internalist, cognitive, and conditionally relativist, correspond respectively to the following functions of his theory: making conceptual connections between value and motivation; gaining knowledge about what is of value; taking value as *non-simpliciter*, but rather as value for different populations. Clearly, this approach to folk moral philosophy in Lewis's value theory abides with Lewis's commitment to dispositional value as conditionally relative, say to a particular population. By analogy, as I argue in **subsection 5.2**, in the Lewisian metaphysics of possible worlds (1986a: 134), although characterized by Lewis as unsystematic, folk theory is described in Humean terms as complying with common sense: as a body of theory, which "we are reasonable to believe. (*Most of it*)."

Additionally, as discussed in **section 3**, Lewis (1972: 170) distinguishes between possible languages or grammars, as abstract semantic systems and symbols with world aspects, and actual languages particular to psychological or sociological descriptions of populations. This distinction comes with recognizing that psychological and social facts are described by the use of an abstract semantic system by persons or populations, which is submitted to the semantical treatment of truth conditions. Putting forth the question of the verification of factual statements, Lewis's own account of radical interpretation (1974a: 333) leaves open

different things: "A man's intention is what he aims at or chooses; his motive is what determines the aim or choice; and I suppose that "determines" must be here another word for 'causes'".

the definition of a 'truth condition':

"[...] in particular, whether it should specify only the actual truth value of the sentence, or whether it should also specify what the truth value would be in various counterfactual situations not too remote from actuality, or whether it should specify the truth value at all possible worlds whatever."

As already mentioned, Lewis is careful not to endow common sense with absolute philosophical authority. Therefore, Lewis (1986a: 134) argues "sometimes common sense may properly be corrected, when the earned credence that is gained by making theory more systematic, more than makes up for the inherited credence that is lost".

I now continue to argue that Lewis's theory of convention, as well as his metaphysical account of modal realism in possible worlds, indicate Lewis's efforts to systematize folk theory, whilst at the same time offering a solution to the problems posed by the anti-Humean accounts of Desire-as-Belief, in their following principles of Lewis's later dispositional theory of value.⁴⁰ I present two philosophical approaches, which, as I show, Lewis adopts in his attempt to analyze and interpret systematically what he terms as folk 'moral science': a conventionalist approach and a modal metaphysical one.

5.1 Lewis's conventionalist approach

In the conventionalist approach, Lewis adapts the Humean view applying game theoretical terms and methods for analyzing problems of coordination.⁴¹ I here argue that Lewis's conventionalist approach is underpinned by the Lewisian possible worlds modal realism, in which Lewis upholds value relativism with regard to worlds. Moreover, I explain Lewis's conventionalist account of analyticity as relative to a population of language users and as determinant of which possible language is the actual language in a population.

⁴⁰ Here, I also take into consideration Plantinga's argument (2003: 223) that, a problem of Lewisian modal realism is its diversion from common opinion, because of the modal reductionism of Lewisian modal realism, despite Lewis's effort to do metaphysics with due respect only to the vocabulary "in which common opinion is commonly expressed".

⁴¹ Putnam's argument for conventionalism (1983: 170), as opposed to naturalism, also explains the occurrence of conventionalism as a theme in analytic philosophy. Putnam (1983: 174-175) refers to the Lewisian game theoretical treatment of conventions as arbitrary, which is opposed to our present view of what is natural.

Lewis (1969: 95-97) describes norms as behavioral, or other, regularities, to which we believe we ought to conform. For this reason, Lewis defines a convention to be always one of several alternative conventions based on preferences.⁴² This definition shows Lewis's early support of a Humean view, which Lewis (1969: 3) broadly situates in Hume's discussion of justice, or justness, and property.⁴³ In this perspective, for Lewis (1969: 93-96), a convention is only a social contract under certain conditions pertaining to the conventionalist notions of mutual preferences and mutual expectations, which are shared as 'common knowledge'.⁴⁴ For instance, conventionalist behavior (Lewis, 1969: 140; Lewis, 1969: 156) can be described as a pattern of agents' expected mutual expectations of expectations of desire of acting on the necessary conditions; or a convention may simply fail to be a social contract and vice versa (Lewis, 1969: 96). I explained that in his dispositional value theory, Lewis examines and rejects different anti-Humean theses to argue for the notions of objective value and ethical reality in folk theoretical terms. In his preceding theory of convention, Lewis's interpretation of Humeanism (1969: 3-4) explores issues around the origin of justice and a general sense of common interest, by using the game theoretical notion of 'coordination-equilibria'.⁴⁵ Lewis (1969: 8) offers

⁴² Lewis (1969: 97) compares and contrasts a convention to a social contract, in Lockean terms, which does not need to have other alternative than the state of nature. Lewis (1969: 97) undertakes this comparison, because convention requires a type of behavior, which he terms as coordination, for achieving what he terms as 'coordination equilibria'. In the social contract, the alternative of the state of nature does not need to be a state, in which 'coordination equilibria' are achieved.

⁴³ I refer here to Hume's metaphysical view of justness in terms of agnosticism towards providence and a future state; see (Hume, 2007: XI§4, 97, XI§10, 98).

⁴⁴ Lewis here adapts the game theoretical thinking articulated in the most commonly applied solution of the Nash equilibrium (Rapoport, 1970; Osborne and Rubinstein, 1994), by introducing the notion of convention, which has its basis on the "common knowledge" of rationality (Lewis, 1969: 70). Conventions need to abide to specific conditions, such as regularity in behavior, but also systems of mutual expectations and mutual preferences, which are shared as common knowledge in a certain population (Lewis, 1969: 58; 78). Lewis describes conventions as adaptable (Lewis, 1969: 75) and substitutable (Lewis, 1969: 85), as long as the 'common knowledge' requirement is satisfied: "[...]a basis for common knowledge generates higher-order expectations with the aid of preexisting higher-order expectations of rationality" (Lewis, 1969: 57).

⁴⁵ This game theoretical notion advances Nash's 'equilibrium', a basic concept for strategic games between rational agents, which has a consequentialist approach; see (Osborne and Rubinstein, 1994: 11-30). According to Osborne and Rubinstein (1994: 84), Lewis introduces the concept of 'common knowledge', to describe an event, "[...] if not only is it mutual knowledge but also each individual knows that all other individuals know it, each individual knows that all other individuals know that all the individuals know it, and so on" (Osborne and Rubinstein, 1994: 73). For Moses (2008: 630), in Lewisian game theory "common knowledge" is taken to be what "any agent X" knows.

Another interpretation by Staudacher (2010: 7) describes the Lewisian conventionalist project, as "one which relates meaning, conventions, and social norms as three independently characterized notions", whereas meaning is taken to be literal meaning. In the conventionalist case, the phenomenon of convention is not homogeneous – Lewis (1969: 12) describes the conventionalist project not primarily as uniform combinations, but as 'coordination equilibria' – and it should not be conflated with "customs", "traditions", or "established practices". For Staudacher (2010: 19), distinguishing between conventions and social norms should be sufficient to understand the

below an exemplified description of problems of coordination encountered when striving for 'equilibria':

"Two or more agents must each choose one of several alternative actions. Often all the agents have the same set of alternative actions, but that is not necessary. The outcomes the agents want to produce or prevent are determined jointly by the actions of all the agents. So the outcome of any action an agent might choose depends on the actions of the other agents. That is why [...] each must choose what to do according to his expectations about what the others will do."

Here, I take Lewis's 'coordination-equilibrium' as a modal notion, which enables Lewis to analyze social behavior in respect to folk theory by using several examples of conventions. Such examples are typical of Lewisian coordination problems, while their descriptions in informal hypotheses adhere to the modal terms of 'possibility' and 'necessity'.⁴⁶

Lewis (1969: 14) argues that some combinations of the agents' actions will result in 'coordination equilibrium', which is "a combination in which no one would have been better off had any one agent alone acted otherwise, either himself or someone else". Following modal reasoning, what Lewis (1969: 8) argues here is that it should not be expected that a 'coordination equilibrium' will produce an outcome that is best for even one of the agents; but that "it is possible that some or all of the agents would have been better off if some or all had acted differently", and that it "is not possible that any one of the agents would have been better off if he alone had acted differently and all the rest had acted just as they did".⁴⁷ Offering

conventionalist project. The latter interpretation poses the problem of analyticity with regard to the conventionalist project, also beyond the Lewisian approach (2000b [1976b]: 144), which raises the question, not of whether one conforms to conventions, but of which conventions one conforms to.

⁴⁶ Lewis (1969: 6-7) offers several examples, amongst which are the following:

- a) A group of campers are looking for firewood. The direction that each individual camper takes is not important, but it is likely that for any two, who will take the same direction, since they will cover the same ground, the one who gets there later will not find any firewood. So, each camper must choose a direction according to his expectations about the others' direction, but each camper must choose a direction different from anyone else's.
- b) A group of hunters are hunting for food in the wilderness. If they hunt separately, each one of them can catch rabbits and eat badly, but if they hunt together, they can catch stags and eat well. If at least one of them deserts the group to hunt rabbits, the stags will get away, and the rest of the hunters will not have any food unless they desert, too. So, each hunter must choose whether they stay with the stag hunt, based on their expectations that no one will desert, or desert themselves.
- c) A group of farmers are deciding on how to divide up some land, so that each of them gets the exclusive use of one plot of land. It is not important who gets to use which plot, as long as each individual farmer never uses the same plot as another farmer and no plot goes to waste. So, each farmer must choose a plot according to his or her expectations about the plots others will use, and the plot, which will be left for himself or herself.

⁴⁷ Lewis (1969: 8) also argues that if a combination is best for everyone, this must be equilibrium. A 'coordination equilibrium' is not to be confused with cases of group imitation:

additional examples of how the aforementioned coordination problems may not reach a 'coordination equilibrium' (Lewis, 1969: 45-48), Lewis (1969: 33-36) explains that a 'coordination equilibrium' differs from a simple 'equilibrium' in that the former involves the common understanding of the coordination problem by means of communicating the problem and coming up with alternative actions, which involve producing mutual expectations by means of agreements or precedents. Lewis (1969: 22) also differentiates between 'coordination equilibrium' and 'proper equilibrium', with the latter being the case that "if each agent likes it better than any other combination he could have reached, given the others' choices".

The Lewisian notion of analyticity relies, as I discussed in **subsection 3.1**, on informal hypothesis and modal reasoning. Although the Lewisian modal realist view is not yet developed in Lewis's theory of convention, possible worlds are discussed as models in conventionalist terms, by first introducing and then applying analyticity as a notion for explaining possible worlds. As Lewis (1969: 207) argues:

"Possible worlds are models, or indices of models, or diagrams of models (state descriptions), or theories of models (complete novels) for some sufficiently rich language. But not all the models for the language, only those that satisfy its analytic truths!"

Quine (cited in Lewis, 1969: xii) emphasizes that, in Lewis's problematization of the analytic-synthetic distinction in his theory of convention, Lewis rests "the notion of analyticity on the notion of possible worlds", which is a hypothetical entity. This is evident in Lewis's consideration of analyticity in conventionalist terms. In his description of the dependency of modal reasoning on facts about possible worlds and the language of any given agent (Lewis, 1969: 207), which, in turn, is dependent

"[...] because of their conditional preference for doing something if the others do, since the regularity that persists by this mutual imitation is not necessarily a convention. For the situation may not be one in which coincidence of interests predominates over conflict. Some sort of equilibrium is sustained, but it may not be a coordination equilibrium" (Lewis, 1969: 120).

Likewise, cases of replication do not indicate a back and forth interaction between people, but, according to Lewis (1969: 32), "[...] a process in which one person works out the consequences of his beliefs about the world – a world he believes to include other people who are working out the consequences of their beliefs, including their belief in other people who [...]". The latter case is reminiscent of the Rawlsian account of value pluralism in democratic societies (Rawls, 1999). However, Lewis (1969: 93-94) interprets the Rawlsian view as an obligation to fair play for reasons of reciprocation of benefits received through others' obedience to a *status quo* situation, which is an attitude expressing a preference to what Lewis calls the *status quo*, rather than the state of nature (Lewis, 1969: 90-91) in Rousseau's terms (1984). Lewis's conventionalist interpretation of equilibrium (1969: 90-95) considers three states: the *status quo*, the state of nature and the lone disobedience; which are ranked and structured according to preferences, to form social contracts, conventions or a combination of the two.

on the conventions the agent is party to, Lewis (1969: 208) must take analyticity generally and necessarily as truth in all possible worlds: "All that human convention can do is to select one verbal expression rather than another to enjoy the privilege of truth by virtue of the facts about the possibility of worlds".⁴⁸ We saw that, for Lewis (1969: 208), "these conventions are regularities in behavior, sustained by an interest in coordination and an expectation that others will do their part".⁴⁹

Raising a Carnapian objection to Quine's view of analyticity in terms of semantics, which analyze truth and analyticity in relation to an agent, or a population of agents (Lewis, 1969: 204), Lewis (1969: 207) provides an account of what he calls "the proper kind of analyticity", which is relative to a population of language users. In this account, Lewis (1969: 206) includes an additional account of any possible language, in such a way that the analytic sentences of a given possible language are identifiable, as well as an account of "the mental, behavioral and cultural factors in a population", which are determinant of the analytic sentences of their language, in turn deciding "which possible language is theirs". Significantly, as Lewis (1969: 207) shows, his conventionalist task could not do away with the notion of possible worlds, which he applies in *Convention* as a semantic model, rather than as a modal metaphysical notion.

⁴⁸ Lewis (1969: 2) goes as far as arguing that language conventions "are a myth" and that language usage simply "conforms to regularities" performed by a population of language users. Later on, Lewis (1984a: 224) argues against a specific semantical account of descriptivism, which he recognizes in Putnam's non-traditional anti-realism and model-theoretical reasoning, and he terms as "global descriptivism". Lewis's defense of his doctrine of realism about possible worlds (1984a: 232) argues against Putnam's semantic and epistemic approach: "The metaphysics of realism survives unscathed. What does suffer, if Putnam has his way, is realist semantics and epistemology".

⁴⁹ This early modal thinking is consistent with Lewis's later view of possible worlds (1986b: 5) as causally and spatiotemporally isolated. This view is exemplified in the Lewisian statement that if everything is held fixed after making one change, a world is not possible at all. As Lewis (1986b: 5) explains:

"We dream of considering a world where the antecedent holds but everything else is just as it actually is, the truth of the antecedent being the one difference between the world and ours. No hope. Differences never come singly, but in infinite multitudes. Take, if you can, a world that differs from ours only in that Caesar did not cross the Rubicon. Are his predicament and ambitions there just as they actually are? The regularities of his character? The psychological laws exemplified by his decision? The orders of the day in his camp? The preparation of the boats? The sound of splashing oars? Hold everything else fixed after making one change, and you will not have a possible world at all."

5.2 Lewis's modal metaphysical approach

I have argued in section 4 that Lewis's value theory departs from the canonical anti-Humean thesis, partly due to its rejection of the recognition of values as values *simpliciter*: that is in terms of metaphysical necessity. Instead, the modal status of the equivalence between value and what one is disposed to value, which is a piece of philosophical analysis, however 'unobviously' analytic and 'unobviously' true (Lewis, 1999: 85), determines the Lewisian analysis of value theory (1999: 85-89), thus preserving the unobvious analyticity of the dispositional value theory. Lewis's modal metaphysical interpretation, in its *de re* and *de dicto* analysis of belief and desire, improves Lewis's dispositional value theory. I argue that Lewis applies *de re* and *de dicto* representations of value, which he uses in representations of beliefs and desires of persons in possible worlds informed by the early modal reasoning traced in his conventionalist approach. I explain that standardizing on properties and on propositions, these representations allow for uniform propositional objects in order to systematize common sense psychology (Lewis, 1979a: 514).

De re and *de dicto* representations correspond respectively to properties and propositions. However, Lewis (1979a: 514-515) argues that we must standardize on properties, taken in the broad sense of natural, intrinsic, but also in virtue of objects relations, on the premise of the functionalism of propositional objects; as Lewis (1979a: 514) argues: "when propositional objects will do, property objects also will do; sometimes property objects will do and propositional objects won't". Although *de dicto* and *de re* analyses do not need the notion of possible worlds, when possible worlds enter the picture, Lewis's proposal to standardize on properties becomes clearer with regard to folk ethics. In the modal realist view, we can self-ascribe properties in correspondence to propositional beliefs, which are also locational properties in logical and ordinary space. In this manner, we also locate or distinguish ourselves as members of populations or subpopulations, whose boundaries may not follow the borders of worlds; in the latter case, self-ascribed properties are objects of beliefs, rather than "believed-true propositions" (Lewis, 1979a: 518-519). This view supports Lewis's preference towards systematizing folk theory in terms of properties rather than propositional objects (Lewis, 1979a: 533-534). Lewis (1979a: 522) takes his preference as far as to propose that self-locating belief must be understood in terms of self-ascription of properties, which is

consistent with Lewis's account of *de se* modality.

Drawing upon both *de re* and *de dicto* modals, Lewis applies the modal *de se* in his value theory to indicate what he terms as 'egocentric modality', which refers both to ascribed and self-ascribed properties and propositions. In *de dicto* and *de re* terms, for Lewis's value theory (2000 [1989a]: 79), a putative value can be *de se*, a 'property', but it can also be *de dicto*, a 'proposition'. In *de se* modality, one refers to a 'property', in one sense of that word, or to an 'egocentric proposition'.⁵⁰ In possible worlds, propositions are objects of thought, which "are supposed to be capable of giving the content of what we know, believe, and desire" (Lewis, 1986a: 55). For Lewis (1986a: 55), because some propositions are egocentric, "irreducibly *de se*", they cannot be captured by propositions, which are true relatively only to worlds. In Lewisian modal realist terms, the objects of thought are possible individuals, which can be possible worlds. An egocentric proposition, which holds for some people, but not for others, can be taken as a property: that is, as a set of individuals (Lewis, 1986a: 54), which can be a set of worlds (Lewis, 1986a: 102) – although, not any set of worlds will give the content of a possible thought (Lewis, 1986a: 106). Arguably, here Lewis (1979a: 534) is consistent with his view that degrees of *de se* belief-desire are attached to regions of properties and not of propositions, rendering the space of possibilities no longer as the space of worlds. Additionally, since, for Lewis (1979a: 542), we rarely know essences, while beliefs *de re* in general are not beliefs (Lewis, 1979a: 538), representation *de re* in accordance with worlds is especially useful. This is because existence according to a world (Lewis, 1986a: 194), in *de re* modality, expresses "the potentiality and essence of things" (Lewis, 1986a: 12) as quantification over possible individuals.

Specifically, Lewis adopts the view of his modal metaphysical counterpart theory, where possible individuals are world-bound, meaning that they are wholly part of one world together with their surroundings (Lewis, 1986a: 214), rather than transworld individuals. As Lewis (1969: 208) argues:

"When have we one possible in two worlds, and when have we two similar possibles in two worlds? A drastic theory settles the question easily: nothing is ever in more than one possible world. This extreme solution shows that the difficulty is not serious, perhaps less drastic solutions would be even better."

Instead of worlds understood in the canonical view as discussed in **section 2**, in

⁵⁰ For an analysis of egocentric logic, see Prior's account (1977) mentioned in Lewis (1973b: 96).

Lewisian modal realism (1986a: 57), set-theoretic constructions out of *possibilia* allowing for the assessment of a possible outcome, under the hypothesis that it is the actual, or the non-actual outcome,⁵¹ can serve for propositions.⁵² In an example taken from the Lewisian philosophy of language of different individual world-bound speakers, who are part of different worlds, their speaker-relativity requires that *possibilia* enter the picture, no matter what the semantic values themselves are and how they are ascribed to the speakers (Lewis, 1986a: 42). Furthermore, Lewis (1986a: 42) argues that *possibilia* may enter into the construction of the semantic values themselves, allowing for modal *possibilist* interpretations. For example, in the case of individual speakers at different worlds, since, by our thought, we ascribe properties to an individual, singular propositions accord with our successful, or unsuccessful, ascribing with regard to their truth value. In Lewis's own words (1986a: 58):

"This much is true: somehow, by our thought, we do ascribe properties to individuals. (Not by thought alone, of course, special cases aside; rather, by thought plus the relations of subject to environment.) Whenever you ascribe a property to an individual, there is the pair of that property and that individual. So your accomplishments in property ascribing can be characterized in terms of property-individual pairs, in other words in terms of the singular propositions that are true according to your ascribing."

Lewis further states that the use of singular propositions for characterizing the speaker's belief and desire in a narrow psychological sense is therefore unsatisfactory. Lewis (1986a: 185) would rather support the view that the analysis of a given world-bound speaker's content of belief, but also the understanding of propositions as contents of belief and desire, are not dependent on the truth value of a subject's content of belief, but on "the psychology of belief-desire" (Lewis, 1986a: 36). By ascribing the same properties to the same things, the same singular

⁵¹ This view is more broadly consistent with modal reasoning, which allows for assessing possible outcomes, actual and non-actual alike; for instance, in counterfactual constructions as combinations of tense and modality, of which a model-theoretic example is the tree-branching time structure (Gamut, 1991: 40-41).

⁵² Lewis (1986a: 4) uses possible worlds terminology to argue that *possibilia* are to be understood as modal notions, which constitute the Lewisian ontology, and which must be taken as "literal truths", despite any ontological costs that such an understanding brings. Here, it can be argued that Lewis does not follow a Humean principle of recombination for individuation, by virtue of the Humean denial "of necessary connections between distinct existences" (Lewis, 1986a: 87), in his statement that: "I cannot altogether accept the formulation: anything can coexist with anything" (Lewis, 1986a: 88). This view is consistent with Lewis's view of transworld individuation; see **Appendix I**, "Objections and Replies", I.

propositions are true according to the speaker and to us (Lewis, 1986a: 58-59).

The above complies with Lewis's treatment of general semantics, as I discussed in **section 3**, which Lewis also applies in his dispositional value theory for systematizing the analysis of the content of belief-desire with respect to speaker and to world. Moreover, I have already discussed that the aforementioned Lewisian treatment of general semantics is consistent with Lewis's early conventionalist approach to the problem of analyticity.⁵³

Returning to the example of the world-bound speaker, 'possibles', or *possibilia*, are introduced in the construction of the semantic values themselves. Instead of an "extreme external" view of the treatment of language,⁵⁴ which Lewis (1986a: 43) describes as: the assignment of semantic values relative to a speaker; the equation of the semantic truth of sentences with mere truth values; fitting of semantic values for modifiers and connectives to make them into functions from and to truth values; Lewis introduces modal *possibilia* to construe semantic values in possible worlds semantics. Lewis (1986a: 44) proposes, for instance, the inclusion of the modifier 'possibly' into a language, to account for the truth value of a sentence if and only if it is true under some shift of world.

To show this, Lewis (1986a: 45) uses again the example of the world-bound speaker arguing that, in Lewisian 'genuine' modal realism, taking that a speaker's world is given and that the speaker is world-bound, when we shift worlds, which are spatiotemporally disconnected concrete entities, in connection with 'possibly', since we do not necessarily shift speakers, the speaker may disappear. For instance, if we shift to a world where there are no speakers at all, or where there is no counterpart of a given speaker, following Lewis's proposition (1986a: 44) of including the modal 'possibly' into a language, the speaker's utterances are not true. Despite starting out being fixed to speakers, Lewis (1986a: 46) points out that worlds vary independently.⁵⁵

⁵³ Lewis (1969: 208) introduces in his theory of convention, the proposition that 'possibles' can be used in systematic ontology, complying with his later explicative and advancing treatment of possibles as *possibilia* in his modal metaphysical theory. As I have mentioned, Lewis distinguishes duplicates from counterparts, whom he construes as world-bound *possibilia*, meaning individuals taken together with their surroundings.

⁵⁴ This treatment of language purports to an externalist view of linguistic theory and the philosophy of language.

⁵⁵ Likewise, as in modal realism, in temporal realism, given a speaker his time is given, but when we shift times in connection, for instance, with the past, we never shift speakers (Lewis, 1986a: 46). The past is considered fixed for Lewis (1979b: 461-462); for instance, as a state of affairs at an actual given time *t* and only under the same range of suppositions.

Lewis (1986a: 28-30) suggests that, although coming out as a special case of egocentric belief, the theoretical term *de se* modality is also useful for analyzing the content of someone's system of knowledge, belief and desire at worlds. Taking the Lewisian view of the world-bound speaker, the egocentric part offers epistemic and doxastic accessibility only to one world. This means that, in *de se* modality, the belief-desire would not be about the world, but about an individual's own place therein, complying with Lewis's view (1986a: 214) that a possible individual is one that is world-bound, meaning that he or she is wholly part of one world. Taken as a proposition, the content of someone's belief-desire can be captured by the believer's doxastic alternatives, which is a class of possible individuals that is also a class of possible worlds (Lewis: 1986a: 28). As Lewis (1986a: 30) states, a doxastic alternative *simpliciter* is "a possible individual, who gets a non-zero share of probability, but the non-zero shares are not all equal". For Lewis (2000: 73-74), the class of doxastic alternatives gives the content of desire: desire *de se*, or egocentric, or essentially indexical desire.⁵⁶

⁵⁶ Lewis's interpretation of *de se* modality adapts the Lewisian modal realist philosophy of counterparts to Lewis's broader metaphysical view of identity. Lewis (1986a: 217) does not completely reject transworld identity; we exist according to many worlds. Different worlds offer *de re* representation of an individual that exists. Moreover, Lewis (1986a: 211) claims that he does not totally deny the existence of transworld individuals; "[...] yet there is a sense in which I say that they cannot possibly exist". For Lewis (1986a: 211), this means that transworld individuals are "impossible individuals", which is a terminological stipulation. Lewis (1986a: 215) solves the problem of transworld identity with his counterpart theory, which is the theory of person-stages: "As for modal contexts, we should note that two possible individuals are counterparts iff there is some possible individual of which they both are stages"; rather than "[...] the theory that ordinary things are transworld individuals unified by counterpart relations among their stages" (Lewis, 1986a: 217). The difference between Lewis's theory of counterparts and the latter theory does not consist in a disagreement about modality, but in an "extensive semantic disagreement": "[...] about which of the things my opponent and I both believe in are rightly called persons, or sticks, or stones" (Lewis, 1986a: 217). See also Lewis on transworld individuation in **Appendix I**, "Objections and Replies", 1.

6

Lewis's value pluralism

In this section, I discuss the influence of two basic Lewisian modal realist positions upon Lewis's philosophical view of pluralism, also taken as value pluralism, in order to explain how these positions inform Lewis's substantive thesis of the plurality of worlds, which Lewis equates with his modal realism. This analysis allows me to present Lewis's value pluralism in respect to his schematic propositions for deontic ethics.

I start from the recognition that, for understanding Lewis's philosophical view of pluralism, we need to consider two basic proposals inherent in Lewisian modal realism: i) taking language as general semantics, yet as applied in a population at a world; ii) not taking modality as quantification over parts of actuality.

As we saw in **sections 1** and **4**, Lewis (1983: xi) articulates his objection to the method, common to analytic philosophy, of taking language "as a starting point in the analysis of thought and modality",⁵⁷ while he proposes to consider language as an integration of "formal semantics into a broader account of the use of language in social interaction"; what, as we saw in **section 5**, Lewis (1969: 195) terms as semantics in a population. For this reason, Lewis supports modal reasoning, which is a mode of reasoning relying, as methods of general reasoning, on informal hypotheses and on mathematical deduction proceeding from axiomatic premises (Lewis, 1986a: 113-115). Furthermore, Lewis (1983a: 22) specifies that, though raising general questions on the use of inductive reasoning, his modal realism evokes skepticism towards induction, because of its potential fallibility in accounting for modal possibilities

⁵⁷ See also **section 1**, including Burge's claim against logical positivism.

when the metaphysical neutrality of reason is taken into consideration.⁵⁸

Construing possibilities in modal reasoning, as we saw in **section 5**, Lewisian modal realism allows for the assessment of a possible outcome, under the hypothesis that it is the actual, or the non-actual outcome,⁵⁹ but also for the use of counterfactual conditionals.⁶⁰ Here, it is important that the terms 'actual' and 'world' are taken as Lewis (1986a: 100) uses them, in Lewisian 'genuine' modal realist terms, which has effects on our understanding of modality and Lewis's thesis of the plurality of worlds. Lewis (1986a: 100) explains that he finds a peculiarity in the ordinary understanding of modality as "quantification over parts of actuality". In Lewis's view (1986a: 100), this is a case where "a plurality of worlds falls victim to analytic actualism" and, consequently, "a plurality of grand worlds does too". Lewis (1986a: 101) further argues that, in adopting such terminological and analytic actualism, or in following the trivial analytic truth that everything is actual, modal realism fails; as Lewis (1986a: 112) points out: "If the other worlds would be just parts of actuality, modal realism would be kaput". Furthermore, Lewis (1986a: 101) explains that "there is no convincing reason to think it analytic that everything is actual".

⁵⁸ In Lewis's argumentation (1983a: 23), inductive reasoning is analogical to unreasonable optimism, because they both consist in ignoring "perfectly good possibilities, which we cannot rule out – of disaster". Lewis (1983a: 22) further argues that as a modal realist, he does not have any reason than anybody else to "give over" his faith in his inductive luck. Instead, he has the same reason that everyone has; as argued by Lewis (1983a: 22-23):

"[...] it is possible, and possible in ever so many ways, that induction will deceive me. That reason is metaphysically neutral. It becomes no better and no worse, if it is reformulated in accordance with one or another ontology of modality".

Lewis (1983b: 25) further explains that, in his modal realism, possibilities can be defined as a subject's "perceptual or doxastic or what-not alternatives", by also clarifying that "[...] possibilities are not always possible worlds".

In Leibnizian philosophy, when following inductive reasoning, probabilistic thinking derives from possibilistic thinking as degree of possibility (Hacking, 2006: 138). Lewis's skepticism towards the use of general inductive reasoning as a method of philosophical analysis derives from his view that induction does not account for possibilities, when taking reason as metaphysically neutral. As Lewis (1986a: 116) again argues: "[...] a modal realist has no right to trust induction – he should turn skeptic forthwith", which, for Lewis (1986a: 117), also accords with common opinion.

⁵⁹ Such an example is offered by means of counterfactual constructions as combinations of tense and modality, of which a model-theoretic interpretation is the tree-branching time structure (Gamut, 1991: 40-41).

⁶⁰ The Lewisian position on construing possibilities in terms of modal, rather than inductive, reasoning is significant for the examination of objections, such as actualist and other positions in metaphysics. In this view, taking the consideration that the title "Counterfactuals" is too narrow for his treatment of the topic of conditionals, Lewis argued that the alternative title "Subjunctive Conditionals" would not have properly delineated the subject of his work. The term "counterfactuals" indicates that they are "for talking about unrealized possibilities", hence we use them when we consider the antecedent to be false – although there is an important use of the counterfactual form, which does not follow this mode of usage, for instance, for not ruling out plausible, yet unwelcome, explanations (Edgington, 2007: 131).

Lewis's critique of the anti-Humean thesis of belief and desire, which I discussed in respect to Lewis's value theory in **section 4**, extends to the realm of doxastic logic, where Lewis also brings into belief the notion of knowledge, which otherwise is not an explicit factor in his value theory. Lewis (1986a: 27) construes doxastically accessible worlds in 'genuine' modal realist terms, where individuals are doxastic alternatives having beliefs, which are dependent on temporal stages, or coexistent as a multiplicity; still, as Lewis (1986a: 29-30) argues, "compartmentalized and fragmented". This case results in someone holding different beliefs in different belief systems, where Lewis (1986a: 31) proposes taking either an intersection, or a union of sets of an individual's different belief systems. By analogy, Lewis (1986a: 31) argues that different doxastic alternatives may exist under two or more belief systems. Significantly, unlike the haecceitist, while understanding propositions as contents of belief and desire, Lewis (1986a: 185) follows *possibilism* to recommend set-theoretic constructions of possible individuals, instead of possible worlds, for characterizing the content of someone's belief system. In the case of subject-predicate, quantified, negative or conjunctive propositions, Lewis (1986a: 185) argues that we need "more complicated set-theoretic constructions out of possible individuals" to serve as contents of knowledge, belief and desire. Furthermore, considering a subject's doxastic alternatives, the content of belief does not adhere uniformly to the truth of any language ascriptions of belief, or to the subject's acceptance of inner sentences, but it is rather attached to what Lewis (1986a: 36) calls "the psychology of belief-desire".

Characteristically for Lewisian modal realism, instrumental rationality is also constitutive of belief, without being either descriptive or normative. Lewis (2005: 316) defines rational decision making according to an agent's belief-desire, which is dependent on what one judges to be right or wrong. Lewis (1986a: 39) explains that, if we attempt to take instrumental rationality as "the only constitutive constraint", indeterminacy between so-called "reasonable" and "perverse" systems will characterize various systems of belief:

"The same behavior that fits a decent reasonable system of belief and desire will also serve countless very peculiar systems. Thus constitutive principles of fit, which impute a measure of instrumental rationality, leave the content of belief radically underdetermined." (Lewis, 1986a: 38)

Hence, for reasons of evaluation, Lewis (1986a: 39) objects to the result of the "radical indeterminacy" between perverse and reasonable systems of belief and desire, solely on the basis of the constitutive constraint of instrumental rationality.⁶¹ Here, Lewis (1996a: 555-556) highlights the significance of his so-called "rule of belief" in "(modal) epistemology" (Lewis, 1996a: 554), which Lewis construes around a system of different rules. Lewis (1996a: 556) defines belief in relation to knowledge loosely, allowing for true belief without knowledge, knowledge without justification, and knowledge without belief.

The advantage of employing *possibilia* to characterize belief content, instead of basing our evaluations on interpretations of the anti-Humean thesis as discussed in section 4, is that inconsistent beliefs may not be ignored. Instead, they can be explained in virtue of their variety, but also, as impossible beliefs, or as impossible beliefs in conflict with one another (Lewis, 1986a: 36). In the *possibilist* interpretation of an individual composed of different temporal stages, which can also be called 'speakers', an individual's knowledge and belief are given by his doxastic and epistemic alternatives; as Lewis (1986a: 41) points out, "these possible momentary individuals, who might, for all he knows or believes, be himself".

The relation of possible worlds accessibility described in Lewis's deontic system of sphere semantics bears additional significance for understanding the varieties of modality, including the deontic modal notions of epistemic and doxastic 'necessity' and 'possibility'.⁶² According to Lewis (1986a: 27), "the content of someone's system of belief about the world is given by his class of doxastically accessible worlds". In capturing the content of someone's belief system, for Lewis (1986a: 28), we do not characterize the content by a class of possible worlds, but by a class of possible individuals: the believer's doxastic alternatives, or possibilities – again, "who might, for all he believes, be himself" (Lewis, 1986a: 29). Drawing from his counterpart theory, Lewis (1986a: 29) argues on the above position that someone believes he has a certain property, if all of his or her doxastic alternatives have that property. A property is here, amongst other things, considered to be a property of inhabiting a world where a certain proposition holds, which is a claim consistent with the Lewisian view of world-bound, instead of transworld, individuals. As Lewis (1986a: 29) claims: "The doxastic alternatives determine the doxastically

⁶¹ For Lewis's treatment of the question of indeterminacy in folk theory, see also **Appendix I**, "Objections and Replies", 2.

⁶² For definitions of the canonical modal terms see (Cocchiarella and Freund, 2008: 253-256).

accessible worlds, though not conversely: a world is accessible, iff at least one of the alternatives inhabit it".

Lewis (1986a: 35) offers an example of accessibility to different doxastic worlds by means of someone holding different beliefs in different systems. If one believes that P by believing P in one system, and also believes that Q, but in another system, then he or she is given doxastically accessible worlds where P holds, but Q doesn't, as well as worlds where Q holds but P doesn't. Lewis calls this process compartmentalization of beliefs. Things become more complicated when P and Q taken jointly imply R, as in every world that is both a P-world and a Q-world is also an R-world, yet someone strongly disbelieves the conjunction of P and Q, as well as disbelieving R. Hence, one fails to believe the consequence of one's two premises taken together, as long as one fails to take them together. Lewis (1986a: 29) adopts a perdurantist view of doxastic alternatives, by claiming that "the same person can have different systems of beliefs at different times". Therefore he or she persists through time by consisting of many different momentary stages at different times, which constitute various person-stages with various systems of belief.

Here, the Lewisian system of spheres presented in *Counterfactuals* meets Lewis's value theory to imply a preliminary and rough approximation of Lewis's dispositional value theory. As Lewis (1986a: 36-37) explains, the introduction of *possibilist* interpretation allows for the interpretation of behavior as a combined system of belief and desire, which admits of degree and takes us to decision theory; for Lewis (1986a: 37): "Saying what it means for behavior to fit a system of degrees of belief and desire is the business of decision theory". Lewis (1986a: 37) further distinguishes between belief and desire:

"On the side of belief, some possible individuals are doxastic alternatives for the subject and others are not. On the side of desire, some individuals belong to the class in which the subject would prefer to be and others do not."

In making this distinction, Lewis (1986a: 38) articulates his critical stance towards the anti-Humean assumption that an assignment of a system of belief and desire to a subject is correct, when the subject's behavioral dispositions fit the system "by serving the assigned desires according to the assigned beliefs". I explained in **section 4** that Lewis argues against this position by undertaking a systematic analysis of the anti-Humean thesis to conclude that desires are not necessarily aligned with suitable beliefs. In possible worlds semantics, this becomes evident in

Lewis's use of doxastic alternatives, aligned with an individual's beliefs, and in his separate use of classes of individuals, aligned with an individual's desires.

Significantly, Lewis (1986a: 2) equates his metaphysical thesis of the plurality of worlds with his 'genuine' modal realism. I suggest that this Lewisian thesis is a substantive thesis, construed on the basis of the Lewisian interrelated notions of pluralism and of inclusivity. Although Lewis (1986a: 2) characterizes differences between worlds only as differences in kind, and not as categorial, he admits that all the worlds, this, or ours, and other, are inclusive things on their own: "The way things are, at its most inclusive, means the way this entire world is" (Lewis, 1986a: 1).⁶³ As we saw previously in **section 2**, Lewisian worlds are inclusive of *possibilia*, which also make for the thesis of the plurality of worlds. In support of modal metaphysical approaches to philosophical analyses that "make reference to possible worlds", or to possible individuals inhabiting possible worlds, Lewis (1986a: 3) departs from the original axiomatic thesis of necessity as truth in all possible worlds, consistently with a canonical interpretation of possible worlds, which I discussed in **section 2**.

Lewis formulates such an example in his system of spheres of possible worlds, using a strict and a variably strict conditional, based on whether the system is universal, as the set of all possible worlds, or not. Lewisian spheres form a dynamic system, in which the counterfactual cannot be defined by the outer modalities, because the truth value of sentences changes in the system, when different spheres are added or deleted, yet without changing the outermost sphere around any world.⁶⁴ Additionally, as we saw in **section 2**, Lewis construes the thesis of the plurality of worlds in his modal realism, by using set theory, specifically the thesis of the plurality of sets, which are concrete, and not abstract, entities.⁶⁵

I argue that Lewis's value theory is relevant to Lewis's thesis of the plurality of worlds, where possible worlds semantics is reconsidered as a universal scheme

⁶³ Here, the notion of inclusivity derives its meaning from a broad ontological approach in Lewisian metaphysics to define any world as a big physical object. Possible worlds, for Lewis (1986a: 1), include: "[...] every stick and every stone"; "you and I"; "the planet Earth, the solar system, the entire Milky Way"; "the remote galaxies we see through telescopes"; "all the bits of empty space between the stars and galaxies"; "anything at any distance"; any world is also inclusive in time, which means having as its parts "long gone primordial clouds of plasma"; "dead dark stars far in the future".

⁶⁴ See figure 2 in (Lewis, 1973b: 18). Here, worlds are taken as individuals in Lewisian terms, which means they are concrete and not abstract. For more, see also Lewis's response to the conflation argument (1986a: 83).

⁶⁵ This move offers, as Lewis (1986a: 4) argues, "[...] an improvement in what Quine calls ideology, paid for in the coin of ontology". Lewis (1986a: 171) defends the same position in his counter analysis of other modal realists, which he characterizes as ersatzists, including linguistic and pictorial ersatz.

with variable values.⁶⁶ Broadly speaking, modal realism works with several restricted modalities, such as nomological, deontic and epistemic. In Lewisian modal realism, the restricting of modalities by accessibility or counterpart relations, like the restriction of quantifiers, is fluid: "inconstant and indeterminate, subject to instant change in response to contextual pressures"; as Lewis (1986a: 8) explains, "not anything goes, but a great deal does". For instance, Lewis's introduction of the variably strict conditional affects the assessment of the truth value of sentences in possible worlds semantics, also in terms of construing Lewis's deontic scheme of sphere semantics.⁶⁷

Taking the notion of the plurality of worlds as a modal thesis in Lewis's metaphysics and in accordance with Lewis's basic modal realist proposals, as outlined at the beginning of this section, Lewis's value theory considered in respect to Lewis's pluralism offers some preliminary conclusions.

1. Lewis places value to propositions, which, consistently with his modal metaphysical thesis of the plurality of worlds, Lewis (2000 [1996b]: 55) conceives as "classes of maximally specific possibilities".⁶⁸ At the same time, Lewis (2000 [1989a]: 89-90) dismisses a merely deontic or nomological necessity of value, as necessity *simpliciter* or so-called metaphysical necessity, by equating value with what we are disposed to value. Here, modal thinking accords with Lewis's equivocally and unobviously analytic theory of value, which I discussed in **section 3**. Introducing the modal of 'necessity' in value theory, as I explained in **section 4**, Lewis's main argument (2000 [1989a]: 90) is that our dispositions to value are necessary and not contingent; for, if the latter was the case, either, strictly speaking, there are no values, or we fall into the error theory requiring that values are all that we want them to be. However, we also saw that, for Lewis (2000 [1989a]: 90), there are no

⁶⁶ Kekes (1993: 9) explains that "pluralism is a theory about the nature of values", which make for "good lives", while its primary concern "[...] is with the relation in which these values stand to each other; the identity of values is of interest to pluralists, *qua* pluralists, only in so far as it is relevant to understanding their relations. Pluralism thus is a theory only about one aspect of good lives. Pluralists may disagree with each other and agree with non-pluralists about the identity of the values that warrant our allegiance". As Kekes (1993: n9) also notes: "In contemporary philosophy, pluralism is used also to denote a metaphysical theory, often referred to as systematic pluralism". Kekes (1993: 10) characterizes pluralism as at once a descriptive and an evaluative theory. Although Lewis's pluralism is applied to his metaphysical notion of possible worlds, which I discuss in this subsection, in my view, Lewis also broadly applies pluralistic principles, such as the conditionality and relativity of value (Kekes, 1993), in his value theory, which I discussed in **subsection 3.1**.

⁶⁷ For a more detailed view of Lewis's system of sphere semantics and the use of conditionals, also in deontic terms, see **Appendix I**, "Objections and Replies", 5.

⁶⁸ Rather than as maximal states of affairs or maximal properties, which is a position held by Lewis's rival modal realist views of 'ersatzism', including Stalnaker, Plantinga, and van Inwagen; see (Lewis, 1973b: 81) and (Berto and Plebani, 2015: 173).

values that we are necessarily disposed to value *per se*, as it becomes clear in our amendment of the dispositional value theory upon introducing the modal 'necessarily'. In this way, Lewis brings forth his view of value pluralism in conformity with his modal thesis of the plurality of worlds, which is construed in Lewisian modal realist deontic terms by a fluid and contextually responsive restriction of modalities.⁶⁹

2. In his value theory, Lewis equally problematizes the notion of a possible valuer or agent, who does not refer to any absolute standpoint, as I showed in **section 4**. Lewis (2000 [1989a]: 82-83) indeed takes that "an absolute version of the 'we' refers to all mankind", in which case a value is a value *simpliciter*, or universal value: meaning, "everyone, always and everywhere, is disposed under ideal conditions to value it". This position acknowledges that all mankind is alike, besides the diversity of valuing due to cultural differences, which, for Lewis, does not provide counterevidence to the recognition of universalism. Nevertheless, as we saw in **section 4**, Lewis (2000 [1989a]: 83) accepts that absolutism does not surpass relativity,⁷⁰ which is a view consistent with the non-abiding to an absolute standpoint of Lewis's system of sphere semantics and counterfactual conditionals.

3. Regarding the accessibility relations to worlds through counterparts determining the nomological accessibility of worlds, which has an effect on Lewis's deontic ethics as I discuss further on, Lewis (1986a: 7) states that "modality is restricted quantification; and restricted from the standpoint of a given world, perhaps ours, by means of so-called accessibility relations". Lewis explains that the applicable restriction may be different from the standpoint of different worlds and thus given by a relation of accessibility. This definition establishes Lewis's view of the nomological accessibility of worlds in modal realist terms – which can also be applied to 'ersatz' worlds, as long as, according to Lewis (1986a: 20), some quantificational analysis is correct.⁷¹ About questions concerning the existence of

⁶⁹ In Lewis's work on counterfactual conditionals in *Counterfactuals*, where Lewis puts forth a theory of deontic ethics using possible worlds semantics to construe his sphere semantics, while giving an axiomatic logic of counterfactuals and defending their underlying basis in comparative similarity of possible worlds. For more, see **Appendix I**, "Objections and Replies", 5.

⁷⁰ For Lewis (2000 [1989a]: 83), 'we' also consists of an indexical, which refers to a population consisting of the valuer or agent and those like him. In this framework, Lewis (1990: 24) also mentions philosophers' use of the indexical 'we', by pointing at the partisanship saturating philosophical discourse: "The lesson is that 'we' may take a philosopher at his word depends crucially on who 'we' are, and what philosophical premises we ourselves argue from".

⁷¹ The Lewisian modal notion of counterparts complies with the view of perdurantism, which Lewis (1986a: 203) favors for temporality: "Perdurance, which I favour for the temporal case, is closer to the counterpart theory which I favour for the modal case; the difference is that counterpart

quantification over possible worlds, which is "commonly restricted by accessibility relations", likewise, "quantification over possible individuals is commonly restricted by counterpart relations"; these accessibility relations by means of counterparts usually involve similarity (Lewis, 1986a: 8). Lewis (1968: 113) explains that standard predicate logic could apply when we quantify over possible worlds, which, as we saw, can also be possible individuals, by using predicates and a domain of quantification suited to the topic of modality. However, we must substitute modal operators by certain expressions, such as new predicates, together with postulates (Lewis, 1968: 115); thus "the domain of quantification is to contain every possible world and everything in every world" (Lewis, 1968: 114). This explanation accounts for Lewis's thesis of the plurality of worlds as plural and inclusive, while it complies with Lewis's modal realist principle of not taking modality as quantification over parts of actuality. Arguably, Lewis derives his explanation from his counterpart theory, which Lewis (1968: 116) recognizes as intertranslatable, despite its rivalry with quantified modal logic.⁷²

At this point, and to summarize the above, we need to pay attention to the Lewisian modal realist position that representation by worlds is the maximal specific *ways* things might be. In this view, possible worlds thinking and modal realism cannot provide causal information, such as explanations about how things, for instance natural phenomena, are caused (Lewis, 1986a: 131). Lewis's modal realism (1986a: 131) only allows us to think that "some worlds have one outcome and some have the other, and that we are who we are, and we are inhabitants of a world with this outcome rather than that"; or, thinking with possible worlds, we may even learn that "the number of the conserved particles has no cause except, of course, for the number of the particles at earlier times". For Lewis, we will not gain information concerning the nomological ways of our world, therefore the modal

theory concentrates on the parts and ignores the transworld individual composed of them". The perdurantist view through time bears an analogy to transworld identity, if we accept that a transworld individual is composed of distinct parts in non-overlapping worlds; for Lewis (1986a: 202), to perdure means to be made up of temporal parts, and for our temporary intrinsics to be properties of these parts, "wherein they differ one from another". Combined with perdurantism, counterpart theory enables Lewis to not completely reject transworld individuals, but to redefine them: "[...] if we quantify over them, then they are part of our world, and we are among them, and we can appropriately call them 'actual'" (Lewis, 1986a: 95). See also **Appendix I**, "Objections and Replies", 1.

⁷² Lewis (1968: 115) clarifies that his counterpart theory is distinct from interpretations of quantified modal logic, provided by others, including Carnap, Kanger, Hintikka, Kripke and Montague, favouring transworld identity, which supports the view that one thing can be in several worlds. The difference of Lewis's counterpart theory lies mainly on the premise that his counterpart relation is not an equivalence relation (Lewis, 1968: 115).

realist position stands against natural teleological approaches, or accounts of divine cosmological creation (Lewis, 1986a: 132).⁷³ As I mentioned in **section 3**, Lewis (1973: 73) takes a revision of a working hypothesis of Ramsey's 1928 theory of lawhood. The Lewisian argument (1973: 74) is that "the laws of a world are the generalizations that fit into the best deductive systems true there", at that world, as well as that "the laws are generalizations which are highly informative about that world". Lewis's emphasis (1973: 75) is placed on ascertaining the character of a world, based on the generalizations that have the status of lawhood at that world, but also on establishing overall similarity and difference between worlds, in drawing upon the similarities and differences of laws at those worlds.

Lewis's account of pluralism, as value pluralism and as a metaphysical thesis of the plurality of worlds, raises obvious questions regarding the application of his value theory, which are not clarified in Lewis's acceptance of contingency (1986a: 7) as far as our world is concerned:

"Contingency plays a part in what is the case in some worlds and what is the case in others; which is dependent upon which world is ours; hence, which are the laws of our world and which worlds are nomologically accessible from ours, hence what is nomologically necessary throughout these worlds."

I would now like to emphasize the importance of the aforementioned considerations for Lewis's value theory and to provide further explanations, by briefly presenting Lewis's application of deontic ethics in his system of sphere semantics.

6.1 Lewis's deontic ethics in sphere semantics

Taking Lewisian sphere semantics, the only nomological concerns abide to accessibility relations to worlds, which are variably restricted based on the standpoint of worlds, and not to causal relations, since worlds are causally and

⁷³ About cosmological and epistemic questions concerning the existence of the universe, including the earth that we inhabit, Lewis (1986a: 132) only appeals to Sciama's extreme form of the "anthropic principle", which he articulates as:

"We ought to find it not remarkable that the physical constants and boundary conditions turn out to permit the evolution of intelligent life, no matter how exceptional the required values may be. For there are many worlds, with all different values of the constants and boundary conditions." For Lewis (1986a: 132), the above principle answers the existential question of why we are here, "[...] by invoking the existence of all conceivable logically self-consistent universes", which is consistent with Lewis's thesis of the plurality of worlds and his 'extreme' or 'genuine' modal realism.

spatio-temporally isolated from each other.⁷⁴ Typically for Lewis's deontic scheme, in Lewis's spheres, accessibility is defined as a relation between worlds, which corresponds to deontic or moral necessity (Lewis, 1973: 5). For Lewis (1973: 8), a morally imperfect world like ours does not belong to its own sphere of accessibility S . This means that, although deontic or moral necessity applies at our imperfect world, it is not always actualized: "sometimes, we do not insist that each world i must belong to its own sphere of accessibility S_i " (Lewis, 1973: 8). Therefore, we assign to each world i as its sphere of accessibility, the set of all morally perfect worlds at which "necessarily ϕ is true at i iff ϕ is true at every morally perfect world" (Lewis, 1973: 8).

In the above example, Lewis (1973: 96) is still working with modality by exploiting formal analogies between modal logic and deontic logic, which is another area of intensional logic, basing a system of spheres on comparative goodness, instead of comparative similarity, of worlds.⁷⁵ The latter systems have

⁷⁴ Lewis (1973b: 24) also describes counterpossibles, which are asserted in argument by reduction, in philosophy, mathematics, and logic, and are thought to be true; however, they are not only false, but impossible, because of their antecedents' denial of "[...] what are thought to be philosophical, mathematical, or even logical truths." These counterpossibles are simply "asserted counterphilosophicals, countermathematicals, and counterlogicals, which only look like examples of vacuously true counterfactuals" (Lewis, 1973b: 24). They are alternative hypotheses, characteristic of some impossible possible worlds, which differ from our world as regards to philosophical and logical truth. It is worth to note that Lewis believes in impossible worlds as theoretical tools, equally useful for philosophical analysis as possible worlds.

⁷⁵ In construing his system of spheres of comparative goodness, Lewis (1973b: 97-98) also encounters and refutes the Stalnakerian Limit Assumption (Lewis, 1973b: 77) that "[...] there never are closer and closer ϕ -worlds to i without end"; also the Stalnakerian Assumption that "[...] there never are two equally close closest ϕ -worlds to i , but rather there is exactly one closest ϕ -world" (Lewis, 1973b: 77). The Lewisian refutation of the Stalnakerian Limit Assumption rests on the premise that it is implausible for spheres based on comparative goodness, because, by reasonable standards of an evaluation based on comparative similarity, "[...] there are respects of difference among worlds that are wholly irrelevant to their comparative goodness" (Lewis, 1973b: 98). In addition, Lewis (1973b: 98) explains that centered, or weakly centered, worlds guarantee one special case of the Limit Assumption, "[...] a smallest ϕ -permitting sphere around i whenever ϕ is true at i ", which is the innermost nonempty sphere around i that is nonempty because it contains i . For Lewis (1973b: 98), this does not prevent "an infinite ascent to better and better worlds" without any innermost sphere to contain best worlds; "[...] for every world, there would be a sphere small enough to exclude it, so the intersection of all nonempty spheres would be empty". When the Stalnakerian Limit Assumption is applied to the Lewisian deontic system of sphere semantics, as Lewis (1973b: 98) argues, it is not plausible:

"Suppose that the goodness of certain worlds increases according to the value of some continuously variable magnitude, and let ϕ be true at all and only those worlds of the proper kind where the value of the magnitude is strictly less than some upper bound. Then there are no best ϕ -worlds."

Lewis (1973b: 81) proposes instead a revision of the Stalnakerian Assumption on the basis that "[...] if we sometimes have more than one closest ϕ -world to i , then we will have more than one admissible selection function. If so, the choice of one of these admissible selection functions is arbitrary." Significantly, Stalnaker's view (1976: 69) of possible worlds rests on the ersatz position that in the Lewisian extreme, as he calls it, modal realism, "[...] the concept of actuality does not distinguish, from an absolute standpoint, the actual world from the others". The introduction of the

different orderings, according to their total net content of moral values, such as, say beauty, truth, and love (Lewis, 1973: 96). In a deontic system, the variably strict conditional in deontic logic plays a role analogical to counterfactuals. Based on constantly strict conditionals, for instance, the conditional 'given that ϕ_i then ψ ', which is dependent on the similarity of spheres $S_{i1}, S_{i2}, S_{i3}, \dots S_{in}$, as sets of worlds, yet is not consistent with any "similarity distance" between one another (Lewis, 1973: 95), Lewis's sphere semantics (1973: 18-19) pose the problem of better worlds and of choice between spheres of accessibility to these worlds. If the counterfactual is taken as a constantly strict conditional, one needs to choose between spheres $S_{i1}, S_{i2}, S_{i3}, \dots S_{in}$ to be the sphere of accessibility to i ; yet there is no correct choice *per se* (Lewis, 1973: 19). In contrast, if counterfactuals are taken as variably strict conditionals, for Lewis, one does not need to choose at all, because the several spheres are all present in S_i together, while each sphere is a non-empty set, which is there to make each counterfactual stage non-vacuously true: S_{i1} for the first stage, S_{i2} for the second stage, and so on. Lewis (1973: 19) suggests that each sphere is right for each stage of the sequence: " S_{i1} was right for the first stage of the sequence, but not the second, S_{i2} was right for the second stage, but not for the first or third, and so on". Such ordering allows for spheres, as stages and as possible worlds, to "coexist in peace" (Lewis, 1973: 19).

Conversely, Lewis (1973: 101) also brings forth the example that, if i is an abnormal world at which there is no preference ordering from the standpoint of i , deontic distinctions collapse. At these worlds, "nothing is obligatory, everything is permissible", which, in deontological terms, means that the outer modalities are vacuous, as well as that "even contradictions are 'necessary'" and "even tautologies are not 'possible'"; therefore there are no reasonable senses of 'necessary' and 'possible' (Lewis, 1973: 101). The comparative possibility operators here express comparative permissibility, which means "comparative goodness-at-best" (Lewis, 1973: 101). In such cases, we might choose to exclude objectionable consequences (Lewis, 1973: 98). Hence, the Lewisian deontic scheme (1973: 98-99) proposes the conditions of normality, universality, and absoluteness, in order to impose a prohibition on abnormal worlds.

absolute standpoint, which Stalnaker (1976: 69) also equates with the standpoint of the actual world, is contrary to the Lewisian dismissal of an absolute standpoint, from which to evaluate, say worlds as best worlds, coming rather from the assumption, in Stalnaker's terms (1976: 69), that the so-called absolute standpoint is a neutral one. For more on the Lewisian use of selection function in possible worlds semantics, see (Garson, 2006: 114-115).

Arguably, the logical conditions of normality, universality and absoluteness, which Lewis proposes for abnormal worlds, correspond to moral values in a supposed Lewisian canonical ethical scheme.⁷⁶ Lewis (1973: 125) in fact defines a canonical interpretation for any *V*-logic, that is a system with a variably strict conditional *V*, "based on a system of spheres satisfying one of the combinations of conditions corresponding to the *V*-logic": a canonical interpretation is an unintended interpretation, whose purpose is mathematical, not metaphysical (Lewis, 1973: 125). Lewis (2000 [1974b]: 19) further recognizes that "our language may be limited in expressive power", when it comes down to modal and conditional operators in deontic systems.

⁷⁶ Lewis (1973b: 103) here reconsiders van Fraassen's approach of values realized at worlds at different degrees, therefore these normative values are neither absolute, nor universal.

7

Lewis's contributions to metaethics, ethics, social and political philosophy

I have so far shown that Lewis's ethics and social philosophy are underpinned by Lewis's defense of the Humean against the anti-Humean thesis, which Lewis applies in his value theory, but also in his formulated accounts of folk 'moral science' and of value pluralism. For this, I have traced the philosophical scope and principles of Lewis's ethics and social philosophy in his seminal works on convention, counterfactuals and the plurality of worlds, by virtue of Lewis's philosophico-logical project of modal realism. Analyzing select remaining examples of Lewis's essays on ethics and social philosophy, for which I offer an overview in **subsection 7.1**, in this section, I present Lewis's contributions and my preliminary brief evaluations of these contributions to canonical accounts of: metaethics in **subsection 7.2**; social and political philosophy in **subsection 7.3**. I return to advance these preliminary evaluations in **section 8**, in view of Lewis's modal realist interpretations.

7.1 An overview of Lewis's publications on ethics and social philosophy

Arguably, Lewis drew to a great extent from his core writings on the philosophy of language, as well as on logic and metaphysics, for his parallel work on ethics and social philosophy, which initially took the form of single papers. I have discussed Lewis's presentation of his systematic metaphysics in *Counterfactuals* and in *The Plurality of Worlds*, by means of the modal notion of possible worlds in Lewis's substantive thesis of the plurality of worlds. As we saw, Lewis (1986a: 2) equates this latter thesis with his modal realism, offering a view of ethics around his own modal

notions of truth and representation in accordance with his modal realist account. Additionally, I have argued that, characteristically for Lewis's conventionalist approach, in *Convention*, Lewis addresses the title's normative ethical term, based on his interpretation of the game theoretical notion of 'coordination equilibria', also drawing upon an initial application of the notion of possible worlds.⁷⁷

In this subsection, I offer an overview of Lewis's publications on ethics and social philosophy, consisting mainly of nineteen essays, collected as previously published papers under the title *Papers in Ethics and Social Philosophy* (2000).⁷⁸ In these essays, Lewis addresses problems in ethics and metaethics responding to their treatment by several moral and political philosophers, including: David Gauthier, Dale Jamieson, Alan McMichael, Robert Nozick, Michael Scriven, Peter Unger; as well as by metaphysicians Peter van Inwagen and Alvin Plantinga. Lewis also comments on the work of John Hawthorne, Herbert Marcuse, John S. Mill, T. M. Scanlon and others. For convenience and for maintaining the coherence of the analysis, I broadly divide the essays into two groups, addressing second-order problems of metaethics and first-order problems of normative ethics with a social and political focus (Copp, 2006: 4-5; Miller, 2003: 1).⁷⁹ As I discuss in **section 8**, Lewis does not dispute normative ethical principles themselves, such as say utilitarianism, but problematizes instead normative interpretations of specific hypothetical cases by offering counterarguments and counter-models.⁸⁰

⁷⁷ Arguably, also with reference to Lewis's notion of his dispositional value theory as a sort of "naturalistic metaphysics" (2000 [1989a]: 68), which I discussed in **section 4**, Lewis's theory was influenced by the fields of behavioural ecology and sociobiology, which incorporated game theoretical and other optimizing models to the study of animal behaviour; see (Allen and Bekoff, 2007: 308) on cognitive ethology.

⁷⁸ Notably, Lewis (2000: 1-4) also includes his own summary of each of the collection's essays. As Lewis (2000: 1) mentions, this collection reprints all previously published papers in ethics and social philosophy, apart from those previously published in the two volumes of the Oxford University Press collection of Lewis's philosophical papers.

⁷⁹ Copp (2006: 3-4) gives the following definitions: "*Normative moral theory* aims to provide answers to the general moral questions that fall into this category"; while: "The second category includes issues or questions *about* morality and moral judgment [...] Answering such questions does not require making a moral claim. It requires making a claim *about* moral claims or *about* morality. This explains why the issues in this category are called "second-order" or "metaethical"". Miller (2003: 1) also explains that metaethics is concerned with questions about questions of normative ethics. Berker (2015: 49-51) has argued that the initial unclear and perhaps epistemically unjustified distinction between ethics and metaethics has been gradually eroded, due to the shift of metaethics "[...] from being devoted exclusively to the meaning of moral terms to also being concerned with the metaphysics, epistemology, and so-on first-order moral claims" (Berker, 2015: 49). As Berker (2015: 49) explains, "[...] many claims put forward as first-order claims within normative ethics are in fact metaethical claims, insofar as they concern the metaphysical dependence of moral matters".

⁸⁰ Knowles (2001: 14-15) argues that there are two ways of offering a systematic treatment of political philosophical problems: either to discuss "[...] foundational theories, enquiring how much they generate a set of principles which can be applied convincingly to a standard list of philosophical problems which our political life throws up"; or, taking problems, which "[...] have

Complementary to the above, Lewis's papers on specific moral philosophical problems in social and political philosophy, such as utilitarianism and consequentialism, as well as toleration and deterrence, include: "Devil's Bargains and the Real World" (1984), "Buy Like a MADman, use like a NUT" (1986), "The Punishment that Leaves Something to Chance" (1987), "The Trap's Dilemma" (1988), "Mill and Milquetoast" (1989), "Academic Appointments: Why Ignore the Advantage of Being Right?" (1989), "Illusory Innocence?" (1996), and "Do We Believe in Penal Substitution?" (1997). I will subsequently analyze select papers on the aforementioned topics in **subsection 7.3**.

Besides Lewis's major papers in ethics and social philosophy, it is worth noting two single essays from Lewis's *Philosophical Papers Vol. II* (1986), included in **subsection 7.3**: the essay "Utilitarianism and Truthfulness" (1972), a single essay with a political philosophical focus, which considers problems of coordination between act-utilitarian agents; and the essay "Are We Free to Break the Laws?" (1981), included in **subsection 7.2**, where Lewis argues that an agent's ability for a certain degree of self-determination, along with the moral disposition to follow contingency, by conforming to a non-traditional view of nomological causality, do not have a deterministic or teleological significance. Finally, outside his collected essays, Lewis defends his own version of modal realism, in response to Blackburn's projectivist and quasi-realist project in "Quasi-Realism is Fictionalism" (2005), which I analyze together with Lewis's previous essay on modal fictionalism, "Truth in Fiction" (1978), in the following **subsection 7.2**.

7.2 Metaethics

Typical of problems in canonical ethics, the following Lewisian examples of thought experiments, which address second-order, metaethical questions in essayistic form, are formulated and explained in Lewisian modal realism. I discuss the essays in the context of relevant discourses outlined in the introductory **section 1**: determinism and non-determinism; compatibilism and incompatibilism; consequentialism; conventionalism; realism, quasi-realism and fictionalism. Here,

given rise to a range of clear answers prior to their theoretical exploration": "We can state the problem in appropriate detail, outline those answers which best encapsulate our intuitions, and seek out a theory from which these results could be derived". Arguably, Lewis follows loosely the second method in order to raise problems with the theory itself.

I only offer brief descriptions of the problems Lewis addresses, along with my preliminary critical evaluations, which I advance in the following **section 8**.

I. In the essays "**Scriven on Human Unpredictability**" and "**Why Ain'cha Rich?**", Lewis analyzes philosophical problems drawing from a combination of probability and game theory.

- In the first essay, Lewis presents an alternative interpretation of Scriven's argument for the impossibility of prediction in decision theory, taking the specific case of a predictor and an avoider. Lewis counter argues Scriven's view of the essential unpredictability of a given agent's actions or behavior. Scriven's paper (1965: 416) defends the main position that it is not possible to achieve a rational, thus correct in his view, prediction of a given agent's behavior, presuming the agent possesses free will. Covering a broad range of examples from the social sciences, Scriven's view derives from critically examining experimental psychological and other data analysis methods, for drawing inferences and correctly predicting individual or group behavior of free will agents. Scriven's introduction of the 'randomness principle' (1965: 417) reinforces his argument against predictive determinism (Scriven, 1965: 425), which further bears sociopolitical insights. For example, regarding group behaviour, Scriven (1965: 422) recognizes that what is assumed to be predictable group behaviour is only the result of the effects of individuals' predictable or unpredictable actions on social groups.

Conversely, Lewis's interpretation (2000 [1966]: 245-246) emphasizes the implicit premise upon which Scriven's main position rests, presupposing ideal conditions accounting for the predictor (Scriven 1965: 423-425); in which case, the avoider, on whom the ideal conditions also apply, will always violate the prediction (Lewis, 2000 [1966]: 246). Lewis (2000 [1966]: 247) brings forth his own notion of the 'compatibility premise' to challenge Scriven's reductionist position that we can generalize starting from accepting that there exists "some possible situation in which choice is in principle unpredictable". In view of his 'compatibility premise', Lewis (2000 [1966]: 247) explains that, for Scriven's argument to uphold, two sets of ideal conditions, which he calls requirement functions, one for each of the agents, must be compatible. This means that any given predictor and avoider must always meet both sets of ideal conditions at once (Lewis 2000 [1966]: 247), which is an ideal in itself, thus probabilistically an impossibility.

- In the second essay "Why Ain'cha Rich?", Lewis revises his analysis of theories of probability taking on Nozick's interpretation of the infamous Newcomb's paradox, which he outlines in the following problem: what are the potential choices regarding two boxes containing assumed amounts of money, which are deposited by a predictor. The Newcomb's paradox poses the question whether the presumed rational choice, that is choosing both boxes, or the presumed irrational choice, that is choosing only one box, is expected to maximize a chooser's gains. Nozick (1969: 139) himself points out that, unlike the other cases of logical paradoxes presented in the same essay, Newcomb's problem is not a clear case at all. Nozick (1969: 134) accounts for a causal relationship between predictor and chooser, considering that there is a case where the chooser has an influence on what the predictor does (Nozick, 1969: 139). Nozick's broader analysis (1969: 121) also accounts for actions probabilistically independent or dependent on states of the world, in which case actions affect, or not, which state obtains (Nozick, 1969: 132), thus introducing an additional to the specific problem perspective by analogy to the possible worlds hypothesis.

Advancing Nozick's position, Lewis (2000 [1981b]: 37) restates Newcomb's problem in his differentiation between two modes of rational or irrational reasoning, which depend on different conceptions of rationality: *V-(ir-)rationality* and *U-(ir-)rationality*. In the first mode of reasoning, the variant conditional *V* is defined as expected utility, which is non-causally determined. In the second mode of reasoning, the variant conditional *U* is defined as expected utility causally determined, but also dependent on credence and value. Lewis (2000 [1981b]: 39) further takes into account a causal relationship between predictor and chooser. However, by reversing Nozick's suggestion, Lewis implies that the predictor has an influence on the chooser. Therefore, for Lewis (2000 [1981b]: 39), the problem is in fact impossible, because the predictor makes the same choice both *V-rational* and *V-irrational*, which is itself an impossibility.⁸¹ Lewis (2000 [1981b]: 40) defends further this position by emphasizing that an agent, who is deliberating on his choice in the revised version of Newcomb's problem, renders the distinction between *V-rational* and *V-irrational* action undefined. As far as the deliberator is concerned, Lewis (2000 [1981b]: 40-41) evaluates the agent to be lacking self-knowledge about the

⁸¹ It can be argued that the problem thus restated is only possible if the chooser has *a priori* knowledge of the predictor's actions; see (Schiffer, 2005).

standards of rationality that he or she is willing to conform with, but also his or her credences and utilities.

As I discussed in **section 4**, Lewis's value theory essays on desire as belief counteract the anti-Humean thesis that contingent desires are necessarily aligned with suitable beliefs, as well as that beliefs are prioritized over desires. Furthermore, Lewis defends a naturalistic, in the sense of being analytic and metaphysical, value theory, as well as conditionally relativist, in his dispositional value theory. I have shown that the essays "Scriven on Human Unpredictability" and "Why Ain'cha Rich?" analyze philosophical problems combining probability with game theory. Lewis criticizes Scriven's reductivist argument for the impossibility of prediction in decision theory, which is based on the false premise of essential and necessary avoidance in response to a prediction. The narrower take of Lewis's interpretation shows that, although, yet also in consistence with, the behaviors of both predictor and avoider, which are themselves predictable, it is not generally possible to make correct predictions about an agent's action or behavior.

Nonetheless, this recognition does not entail that a given agent's actions are essentially unpredictable. By applying approximations to Scriven's ideal sets of conditions and using the 'compatibility premise' as a variant, in the specific case of predictor and avoider we can predict two different outcomes of the given agents' behavior (Lewis, 2000 [1966]: 248). In the Nozick example, Lewis situates the problem of free will in respect to the canonical discourses of determinism and non-determinism, concluding that a given agent's ability for a certain degree of self-determination, taken with an agent's moral disposition to follow contingency, do not bear a deterministic or teleological significance, despite their being subject to a certain extent to causality. Lewis's use of the notion of compatibilism, by reference to the ideal of the 'compatibility premise', in the Scriven example supports a soft-deterministic position against Scriven's non-determinism with regard to predicting agency. Lewis's restating of Newcomb's paradox only shows that there is no conclusive prediction about whether the rational or the irrational choice will maximize one's gains, thus supporting a non-deterministic position, although acknowledging causal relations between the given agents.

In my evaluation, by situating the problems in the canonical accounts of determinism and non-determinism, but also compatibilism and incompatibilism, and offering alternative interpretations, the aforementioned arguments reinforce

Lewis's value theory positions of the conditional relativity of value, but also of the Humean view against the anti-Humean prioritization of belief over desire, or the necessary alignment of suitable beliefs with contingent, but corresponding, desires.

2. Continuing with the above line of inquiry, in "**Are We Free to Break the Laws?**", Lewis brings forth the problem of nomological causality, using the hypothetical example of an agent raising his arm and throwing a stone in a certain direction. In Lewis's description, if an agent did so, the stone would hit a window and the window would be broken, therefore the agent is able to break a window. He is able to do something, such that, if he did so, his act would cause a window to break, or a window-breaking event. Now, Lewis requests that we suppose that the agent were able to throw the stone at great speed, so that it would fly faster than light, or that his arm would move faster than light in the course of the throw. He would be able to do something, such that, if he did so, a law would be broken, or a law-breaking event would take place. Would his act of raising his arm cause a law-breaking event, or is it even a law-breaking event in itself? The above can be formulated in the following two Lewisian claims: i) the agent is able to do something, for instance raising his arm, such that, if he did so, a law would be broken; ii) he is able to break a law. However, taking the two claims together, Lewis asks whether we can conclude that the agent is indeed able to break a law.

To answer the above, Lewis examines the normative accounts of incompatibilist and soft-determinist positions, which uphold both theses in varying degrees of consequentialist and non-consequentialist approaches, together with theories of nomological causality. Although incompatibilists do not recognize a direct causal connection between the raising of one's arm and a law-breaking event, they acknowledge a causal dependency between the two. This position is consistent with soft-determinism, which, as Lewis (1986e [1981]: 295) also recognizes, in order to be upheld, one need not claim incredible law-breaking powers. In this example, Lewis (1986e [1981]: 291) upholds a compatibilist and non-determinist position by demonstrating that the agent is not able to break a law.⁸² In case a law is broken, it is not caused by someone's action, because the course of events has already diverged a little before the action took place, which means that a law has

⁸² In this case, the Lewisian non-deterministic position can also be viewed in game theoretical terms, as an interpretation of games of mixed strategy on the Nash equilibrium, where agents' choices, although regulated by probabilistic rules, are non-deterministic (Osborne and Rubinstein, 1994: 31), while they deliberately introduce randomness into their behaviour (Osborne and Rubinstein, 1994: 37).

already been broken. Neither is it necessary that, if one had not raised his arm, the 'divergence' would not have happened; although, as Lewis (1986e [1981]: 296) argues, "what is true at most, is that if I had not raised my hand, then the first ['divergence'] miracle might not have taken place".⁸³ One's own action has not caused a 'divergence', neither it is necessarily causally dependent on a 'divergence'. The agent is not able to break a law, despite the occurrence of a law-breaking event. Lewis here argues against accounts of nomological causality between an event and perceived consequences of the event. This argument is supported by Lewis's distinction between counterfactual and causal dependence from what he terms as nomic dependence (Lewis, 1973a: 563). For Lewis (1973a: 566-567), a deterministic explanation would hold the view that 'divergence' requires the violation of natural laws. Lewis (1986a: 103) further puts forth an argument, which can be taken as a refinement of the deterministic explanation of 'divergence', that there is some natural break in possible space-times, which is, however, "sufficiently salient within the mathematical universe not to be ad hoc".

In my evaluation, by examining issues of nomological causality in view of accounts of determinism and non-determinism, but also compatibilism and incompatibilism, in this example of applying a scientific parable, Lewis introduces his critique of consequentialism, with which he is also concerned in several problems of social and political philosophy.

3. In his essays "**A Problem About Permission**" and "**Reply to McMichael**", Lewis presents more paradoxes using game theoretical and possible worlds analysis of deontic problems.

- In the first essay, assuming a language game between three players: a master, a slave and a so-called *kibitzer*; and taking as the beginning of the game the set of all worlds, Lewis admits no absolute truth value to the imperative and permissive utterances used by the master. The term 'truth value' is here used as a theoretical term, in Ramsey's sense of a term introducing scientific theory in relation to the theory itself (Lewis, 2000 [1975b]: 25). Following his account of sphere semantics, Lewis (2000 [1975b]: 21-22) argues that this is because the spheres of worlds adjust automatically to make the utterances truthful, but also because, when neither commanding nor permitting, the master and other players use utterances that are truthful at the worlds conforming to their

⁸³ Here, Lewis uses the term "miracle" in the Humean sense of a violation of natural laws; see (Hume, 2007: 83).

beliefs (Lewis, 2000 [1975b]: 22). In view of the above, Lewis (2000 [1975b]: 33) simply states the impossibility of the problem of a definite and correct principle governing an assumed evolution of permissibility of spheres. Lewis (2000 [1975b]: 23) also states that, in this essay, he treats commanding and permitting as part of social practices, which enable the control of a subject by another.

- In the second essay "Reply to McMichael", Lewis argues against an ordering of worlds on the non-commonsensical principle of 'radical utilitarianism', by the application of his possible worlds semantics for deontic conditionals as presented in Lewis's *Counterfactuals*. Stressing the neutrality of deontic conditionals themselves, which renders them applicable to formulating any ethical doctrine (Lewis, 2000 [1978]: 34), Lewis (2000 [1978]: 34) argues that a 'radical utilitarian', who is a good example of a consequentialist, would order worlds on 'radical utilitarian' principles to apply his semantics for deontic conditionals. Nevertheless, as Lewis (2000 [1978]: 35) points out, although "the radical utilitarian should find them much to his liking", "most of us would indeed find these results strange and unacceptable", because of their disagreement with our ordinary moral thought. Lewis also emphasizes that the consequences of a 'radical utilitarian' application of deontic conditionals would be precisely those that McMichael wants to avoid: that is the lack of ethical requirements and prohibitions (Lewis, 2000 [1978]: 34-35).

To illustrate his argument, Lewis (2000 [1978]: 35-36) refers to McMichael's folk example of a bank robbery (84: 1978). In this example, for every world w in which someone robs a bank, there is a better world w_I with overabundance of goodness where someone robs a bank, yet it is permissible that the robber does not confess and return the stolen goods, because of the goodness in that world, which counterbalances the sin. Lewis (2000 [1978]: 36) argues against the appeal of the 'radical utilitarian' to quantifying the overall goodness, or sinfulness, of a possible world, for the ethical assessment of one's own actions, but also for setting obligations and prohibitions. Lewis (2000 [1978]: 36) proposes instead that, "the only relevant good, sinlessness, is not 'a good which may exist in amounts of any size'", which means it is not quantifiable. Confronting the 'radical utilitarian' view with ordinary morality, Lewis also concludes on the relativity of ordinary moral

opinions, which are settled, if ever, on a rare stable equilibrium between the utilitarian and the theological ordering view.

Notably, also concerning the above example, in his essay "Semantic Analyses for Dyadic Deontic Logic" (1974b), Lewis (2000 [1974b]: 15) advances his work on counterfactual conditionals, in an effort to systematize further his deontic system by extracting results for deontic logic and providing clarifications for the use of conditionals in his system of sphere semantics. Lewis (2000 [1974b]: 15) argues that frames, including trivial, normal, universal, limited, separative, closed or linear, are such, if and only if every value structure that they assign is also so. Moreover, Lewis (2000 [1974b]: 15) argues that there is equivalence between two frames, if and only if any interpretation based on either one is also based on the other. Using set-theoretical semantics, in this essay Lewis continues his previous work on deontic logic in *Counterfactuals*, to argue for the possibility of infinite sequences of better worlds in the necessity to construe more than limited frames.

Furthermore, taking an imperfect formal analogy between two different subjects, counterfactual conditionals and dyadic deontic logic, Lewis applies set-theoretical interpretations on the following deontic folk case, which is a generalization of McMichael's example:

- (1) It ought not to be that one is robbed (absolute non-permissibility)
- (2) It ought not to be that one is robbed and then helped (absolute non-permissibility)
- (3) One ought to be helped, if one has been robbed (conditional permissibility)

Since the best of the best possibilities have already been excluded in the case of (1) and (2), deontic principle (3) makes possible the actualization of the best amongst those remaining possibilities. Drawing from four treatments of semantic analyses for dyadic deontic logic, including his own, which deal with the above described case,⁸⁴ Lewis's method (2000 [1974b]: 7) consists in using truth conditionals for formulating value structures, bearing "information about comparisons or gradations of value" for evaluating possible worlds. A deontic set-theoretic interpretation is based *at* a world, *on* a value structure, applying the truth conditional 'iff' for every deontic sentence formalized as $O(A/B)$, where O is the obligation deontic operator and A and B are sentences, or deontic sentences for iteration, at that world and under that interpretation, for defining the dependence

⁸⁴ The other semantic interpretations are by Bengt Hansson, Dagfin Follesdal and Risto Hilpinen, and Bas van Fraassen.

on the evaluations represented by the value structures (Lewis, 2000 [1974b]: 7-8). Lewisian value structures over a set I comprise all four kinds of the following value structures: a selection function f over a set I assigning to subsets a subset on the premise of two truth conditions; as well as a ranking, a nesting and an indirect ranking over I (Lewis, 2000 [1974b]: 8-9).

Lewis also distinguishes between what he calls *trivial* value structures of all four kinds, in which every world is 'inevaluable'; that is a world where the usual truth conditionals of two-valued logic do not apply: where $0=1$, and given $0=1$, "ought nothing or everything to be the case?"; or, in deontic terms, "is everything or nothing permissible?" (Lewis, 2000 [1974b]: 9) Lewis (2000 [1974b]: 10) proposes to ignore trivial value structures and use only the remaining non-trivial, or *normal* value structures, but also *universal* value structures with no 'inevaluable' worlds. Lewis (2000 [1974b]: 10-11) considers what he calls limited value structures, which preclude infinite ascending sequences of better and better worlds, as well as 'separative' value structures, in which any world surpassing various of its rivals separately, also surpasses all of them when taken together. Finally, Lewis (2000 [1974b]: 11-12) considers closed value structures for nestings and linear value structures for indirect rankings, where closure and linearity have no semantic effect. For not excluding the serious possibility of infinite sequences of better worlds, Lewis (2000 [1974b]: 19) concludes that there is the necessity to construe more than limited frames, taking language operators with the truth conditions of all four value structures, permitting iteration, as well as using 'separative' closed nesting frames and ranking frames, with normality, universality and absoluteness as options, for the semantic apparatus. Lewis (2000 [1974b]: 19) further gives three versions, partial choice function frames, nesting frames, and ranking frames.⁸⁵

⁸⁵ Moreover, the relation of equivalence and the condition of validity are used in Lewis's dyadic deontic logic. The first, which is reflexive, symmetric and transitive, defines relations between value structures of all four kinds; for instance, for two value structures to be equivalent, they must be value structures over the same set: if a value structure is "trivial, normal, universal, limited or separative", its equivalent must also be so (Lewis, 2000 [1974b]: 12). The second is a truth condition on a sentence, which validates the sentence under a particular interpretation over a set of worlds I , "iff it is true at every world in I " (Lewis, 2000 [1974b]: 15).

To generalize further his evaluations represented by the value structures, Lewis (2000 [1974b]: 14) points out that evaluations must be carried out between three levels of generality, limited, separative and unrestricted, rather than between the four value structures. Once this has been decided, a class of value structures must be chosen, which, for Lewis (2000 [1974b]: 14), comes down to a preference for this class than for another. Finally, for assigning value structures to worlds, Lewis (2000 [1974b]: 14) also constructs frames, which are families of value structures. The relation of equivalence also applies to frames under certain conditions (Lewis, 2000 [1974b]: 15). Lewis (2000 [1974b]: 15) explains that frames, including trivial, normal, universal, limited, separative, closed or linear, are such, iff every value structure that they assign is also so; as well as that there is

In "A Problem About Permission", Lewis (2000 [1978]: 29-33) describes command and permissibility in game theoretical terms, to reiterate the position that, in possible worlds semantics, what he terms as the 'principle of evolution' of the sphere of permissibility, simply does not solve the problem of permissibility regarding worlds. This is because, although a principle is offered, in Lewis's view: firstly, it is not in itself both definite and clearly correct; and secondly, it is not explained what is the comparative relation of permissibility or near-permissibility at every stage, and if there is evolution, how this relation evolves. Likewise, as we saw, in "Response to McMichael" Lewis argues against an ordering of all worlds on 'radical utilitarian' principles. This is because, aside from such ordering not bearing the optimal utilitarian result, which McMichael advocates, which is maximizing goodness across worlds, maximization is not possible based on the neutralizing principle of "the total balance of good and evil" (Lewis, 2000 [1978]: 34). As Lewis (2000 [1978]: 34-35) points out, a 'radical utilitarian' ordering on the premise of "only the total matters" produces the results of little obligation and much permissibility.

In my evaluation, in both the aforementioned ethics and social philosophy essays, Lewis shows that, neither offering an evolutionary account as such of permission and obligation, nor taking the view of quantifying across all worlds the totality of goodness and evil, based on a neutralizing principle, solve deontic problems at worlds. Lewis remains here consistent with his elaborate analysis of the assessment of value structures, based on the relation of equivalence and the condition of validity, in "Semantic Analyses for Dyadic Deontic Logic", thus defending the view of the serious possibility of better and better worlds. As articulated in his claim about the impossibility of quantifying 'goodness' across worlds, but only taking the principle of 'comparative goodness', or 'comparative goodness-at-best', in his thesis on the plurality of worlds (Lewis, 1986a: 128), Lewis here is also consistent with his non-consequentialist view, which I discuss in the following **subsection 7.3**.

4. In "Evil for Freedom's Sake?", after examining free-will theodicy as a philosophical explanation to the problem of evil (Lewis, 2000 [1993]: 126-127), Lewis (2000 [1993]: 107) argues that selective freedom is a paradoxical concept. In support of this interpretation, Lewis revisits Newcomb's problem to pose the question of

equivalence between two frames, iff any interpretation based on either one is also based on the other.

freedom in compatibilist and incompatibilist terms, only to dismiss the usefulness of the theistic question of selective freedom and to claim agnosticism.

Despite the overall very limited theological allusions, specifically direct references to God, which can be explained by Lewis's self-proclaimed atheism (2000 [1989a]: 159-160; 2000[1993]: 206), as we saw in **subsection 6.1**, Lewis (1973b: 101) states that the preference ordering in what Lewis calls an abnormal world is lacking, because of the inexistence of a god, who would promulgate laws determining the preference ordering from the standpoint of say world *i*. Moreover, Lewis (1973b: 99) argues that the ordering of worlds will differ from their unique standpoints, because, if there are gods promulgating laws, different worlds will be ruled by different gods, who promulgate different laws.

In my evaluation, also in view of Lewis's pluralism, this Lewisian argument challenges the condition of absoluteness, which Lewis suggests it is imposed on a system of spheres, presuming the existence of God in the Christian tradition. Still Lewis (1973b: 99) acknowledges that this condition "is correct for some preference orderings and wrong for others",⁸⁶ which stresses again Lewis's pluralist thesis.

5. In "**Convention: Reply to Jamieson**" and "**Meaning Without Use: Reply to Hawthorne**", Lewis addresses the question of conventions with regard to value relevant to populations and language usage by populations.

- In the first essay, Lewis (2000 [1976b]: 136) argues for and against conventionalist and non-conventionalist views of value theory, by presenting examples of considered conventions, to argue that their analysis is derived from a definition of conventionality as non-conscious, yet relative to populations, and based on a distinction between conditional value preferences and conditionals about value preferences. Jamieson (1975: 81) disputes the ordinariness of,⁸⁷ or alignment with common sense, of Lewis's conception of convention. Lewis (2000 [1976b]: 138) explains that, in his own conventionalist terms, conventions are defined by populations' expectations and value preferences for certain conventions to prevail. Such conventions can be unused, not honored, or outdated. Populations may also expect and prefer that

⁸⁶ Lewis's sparse references to god-like perspectives, as well as his approach to the problem of theodicy, a term drawn from Leibnizian philosophy – see (Jolley, 2005) – can be interpreted as typical of the secularization of metaphysics in the twentieth century.

⁸⁷ Lewis (2000 [1976b]: 144) also distinguishes between conventions, as regularities of behaviour, adopted by mutual expectations and preferences and conventional behaviour, which may not conform to our ordinary views, because, for instance, someone's behaviour may be much more conventional than anybody else's.

established conventions are replaced by alternatives. For Lewis (2000 [1976b]: 139), his definition does not rely on ascertaining the so-called 'unconscious' states of populations, but "need only rely on preferences and expectations that stand ready to manifest themselves as soon as a relevant choice or question arises". Lewis (2000 [1976b]: 141) also clarifies that his concept of 'common knowledge' in a convention relevant to a population need not require "knowledge about hypothetical situations as a condition for something to be a convention", but rather "knowledge of conditional preferences" and certainly not "conditionals about preferences" (Lewis, 2000 [1976b]: 142).

- Continuing with the question of conventions, in **"Meaning Without Use: Reply to Hawthorne"**, Lewis addresses the problem of actual and possible languages, which we saw in respect to Lewis's folk theory. Lewis here defends his conventionalist account, which he originally formulated in his previous essay "Languages and Language" (1983e [1975a]). Lewis's main argument responding to Hawthorne is that we have trust in a language, when the subjective probability of uttering used or unused sentences, which are truthful in this language and its corresponding world, is more than zero. Lewis here problematizes Hawthorne's challenge (1990: 116) that it is not possible to have a purely conventionalist account of language, because trust and truthfulness would be in some fragment of some particular language, rather than in the specific language itself – or, following Lewis (2000 [1992]: 149-150), in every fragment of this language, therefore the fragment will not determine the rules for that language (Hawthorne, 1990: 116).

Starting from the viewpoint that the use of language determines its meaning (Lewis, 2000 [1992]: 145), Lewis brings forth the example of unused languages to stress the validity of the conventionalist account for both used and unused utterances, claiming, contrary to Hawthorne (1990: 117), that unused sentences will get more than zero probability of being uttered (Lewis, 2000 [1992]: 147-148). Hence, it is precisely the truthfulness of the unused sentences, which, for Lewis, will determine in which language we can have trust, should we have to choose between two languages, say L₁ and L₂. The conventionalist aspect of Lewis's argument, as Lewis (1983e [1975a]: 184) previously showed, is not based on the implication of the regularity of trust and truthfulness *simpliciter* in language, but the regularity of trust and truthfulness in some particular language, which follows certain conventions.

Lewis (1983e [1975a]: 179) also recognized that the beliefs and desires of a population are "such as to comprise a fully determinate convention of truthfulness and trust in some definite language", meaning that beliefs in a population are held in acceptance of certain suitable sentences in their language. Hence, we say we have trust in this or that language. Nevertheless, Lewis (1983e [1975a]: 182) clarifies that belief should not be analyzed in terms of acceptance of sentences.

In my evaluation, the above Lewisian interpretations are consistent with Lewis's dispositional value theory, where Lewis defends a conditionally relativist view of value theory, relevant to populations, and established by approximation to canonical methods of defining value. In addition, in using probability theory and possible worlds interpretation, by arguing for the subjective probability of uttering used or unused sentences, which are truthful in a specific language and its corresponding world, to be more than zero, Lewis defends his own conventionalist account in relation to his dispositional value theory, which is applied in his folk theory.

6. In "**Quasi-realism is Fictionalism**", Lewis advances his analysis of the problem of language usage in a population in his discussion of dispositional value, responding to the metaethical discourses of realism, quasi-realism and moral fictionalism.

We saw that Lewis distinguishes between actual languages, which means languages used in a population, and possible, which means unused or abstract, languages. Despite Lewis's admittance to a certain degree of indeterminacy in actual languages, his conventionalist approach supports a high degree of trust and truthfulness in a particular language, which is not decided empirically or contextually, in taking the language's both actual and possible utterances. Blackburn (1984: 264) also distinguishes between a possible and an actual language, by characterizing the former with abstraction, which renders it a possible language, but not necessarily a spoken language. For Blackburn (1984: 264), abstraction is ensured, because nothing empirical or contingent enters our specification of a language: "that will arise when we fit the abstract language we have specified to any particular people, for it is then that the question arises whether they actually use any signs in the ways stipulated". As Blackburn (1984: 264) further argues: "Pure semantics, in this sense, is entirely stipulative – merely a matter of armchair invention". Blackburn agrees with Lewis's interpretation of appointing a certain

degree of indeterminacy to semantics in a population, as an approximation to the canonical method. However, Blackburn (1984: 279) places emphasis in actual languages, bearing semantic interpretations in correspondence with physical facts, which may be determined empirically or contextually – for instance, radical interpretation is partly “about how semantic interpretations sit on top of physical facts about things” – whilst dismissing the usefulness of abstract languages in verificationist terms, as far as a population’s language usage is concerned, because such language cannot be empirically or contextually determined.

Significantly, Lewis (1986a: 23) makes an argument for a case of adjustment, rather than for evaluation of a given belief-desire system’s correctness,⁸⁸ when considering empirical or contextual data, such as perceptual experience – say what is going on in a given subject’s or population’s environment – for evaluating the content of the subject’s or population’s belief-desire as largely correct. This argument is consistent with Lewis’s view of a given subject’s or population’s beliefs and desires comprising of a convention of trust and truthfulness in some definite language, rather than say being empirically determined. Lewis supports the view that the subject’s behavioral dispositions will fit any system of belief and desire, yet this adjustment does not make the given system correct. As Lewis (1986a: 38) argues: “What makes an assignment of a system of belief and desire to a subject correct cannot just be that his behavior and behavioral dispositions fit it by serving the assigned desires and assigned beliefs”. This criticism comes from Lewis’s recognition (2005: 316) that belief does not automatically motivate, rather “one inevitably desires what he judges to be right or good” – likewise, in the case of rationality.

In response to Blackburn’s account defending quasi-realism against fictionalism, Lewis (2005: 318-319) explains further that cases of adjustment are not to be confused with a quasi-realist position, which would imply that one declares that there are no values, while it is nonetheless legitimate, rather than convenient, to continue our discussion about values by using the term value for non-deserving claimants (Lewis, 2000 [1989a]: 93). Such a quasi-realist position, for Lewis (2000 [1989a]: 93), should not be equated with Blackburn’s own quasi-realism, in which Lewis (2005: 315) does not recognize the linguistic dispositions of the realist to describe as “a variety of moral fictionalism”, where we retain morality with

⁸⁸ See also **Appendix I**, “Objections and Replies”, I, on Lewis’s view of the Humean supervenience thesis.

recognizable values, but we treat it as fiction. If Lewis treats morality itself as erroneous, for Blackburn (2005: 326), either one has to abandon it altogether: as in Lewis's definition of a sort of quasi-realism (2000 [1989a]: 93) and in his analysis of Mackie's position (2000 [1989a]: 90); or one must correct the error and retain the non-erroneous part, if the error is found in the first-order discourse practices; or one must not disguise the truth, but also not engage with say "bad theory", if the error is philosophical theoretical (Blackburn, 2005: 331-332). Here Blackburn's disagreement (2005: 327) about identifying moral fictionalism with his quasi-realist thesis is supported by his argument that, in moral fictionalism, one does not do any of the above, just continues treating morality as fiction, that is as make-believing; for instance, by echoing the realist, but without the realist's mistakes (Blackburn, 2005: 327).

Worthy of mentioning is quasi-realism's upholding of projectivism, which, in Lewis's interpretation (2005: 316), means that our moral attitudes "originate within us as a result of contingent aspects of our psychology and upbringing". Conversely, Blackburn (1984: 187) states that:

"The argument for a projective moral theory is in effect a development of the simple thought that moral properties must be given an intelligible connection to the natural ones upon which they somehow depend. It generates a metaphysical motive for projectionism."

When applied in folk ethics, Blackburn's projectivist view (2005: 334) is sceptical concerning any applications of the 'folk-concept of the good', "since there is unlikely to be any property that makes all of them true". Thus, the quasi-realist remains unaffected by folk moral psychology, to side with a Kantian view of the autonomy of ethics from empirical or other external rewards. Blackburn (2005: 335) makes a final argument about the metaphysics adhering to the possible worlds hypothesis, highlighting the usefulness of possible worlds as a model of analysis, rather than as a sort of verificationism for fiction or conviction, about the existence of possible worlds and moral assertions derived from therein.⁸⁹

Lewis's defense of the view (2005: 316-317) that moral values are collective and contingent to worlds brings forth Lewis's argument, against Blackburn's naturalistic position, that moral diversity is not dependent on individual

⁸⁹ Although addressed by various mathematical realists, such as (Chihara, 1998), specifically also in relation to the Lewisian modal realist interpretation, the epistemic question about the existence of possible worlds is not explicitly treated by Lewis in verificationist terms, as we saw in **section 2**.

underlying dispositions, but on the actualization of different, amongst shared, dispositions by different individuals. In emphasis of this view, Lewis refers to his previous essay on fictionalism, where Lewis uses an analysis of folk fiction to reveal its origins in the collective belief worlds of a community. Lewis (1978: 42) counteracts the argument that counterfactual reasoning, as used in fiction, equates with modal informal reasoning of common sense modal opinions in folk ethics, yet without describing the latter.

In my evaluation, Lewis's verificationist argument relies on taking the reliability in terms of truth value of any given language, on the basis of both used and unused utterances in a population of language users. Furthermore, Lewis's conventionalist interpretation of value structures and language usage, which is articulated in his dispositional value theory, problematizes, but without explicitly dismissing, the realism of moral values, and certainly without resorting to fictionalism. The above has obvious implications for Lewis's folk theory, especially his folk 'moral science', which, as I showed in **section 5**, forms a significant aspect of Lewis's ethics and social philosophy. On the other hand, Blackburn's quasi-realism separates questions of language and language usage from questions about the ontology of moral values, which he construes as a case of projectivism of moral properties, rather than as fictionalism. For this reason, folk moral psychology, which, for Blackburn, rests primarily on questions of the use of language, is not to be considered in an account of ethics.

7.3 Ethics, social and political philosophy

Moving onto first-order problems of normative ethics and social and political philosophy, in his sustained critique of utilitarianism, Lewis endorses at the same time a non-consequentialist position; as Lewis (2000 [1984b]: 214) states, "consequentialism is all wrong as everyday ethics, right as a limiting case". For Lewis, consequences matter morally, yet they can be outweighed by other moral concerns. Lewis argues that, in moral terms, we speak of good and evil ways to act, whatever one's actual ends might be, which is a consequentialist concern. Hence, we may face the paradoxical problem that it may be a good act to form an intention, which would be an evil act to fulfill:

"Suppose your ends are morally good ones, so that it would be morally right

to pursue them in an instrumentally rational way. Suppose also that they are urgent enough that it would be morally wrong to pursue them in an instrumentally irrational way. This may be so – let the stakes be high. Then it doesn't matter if we speak of right or wrong in an instrumental or in a moral sense. The two senses coincide, the two paradoxes coalesce." (Lewis, 2000 [1984b]: 202)

As Nolan (2005: 187) argues, here Lewis means that consequentialism is only right in the extreme case of deterrence of nuclear catastrophe, which I discuss further on in this subsection. By virtue of Lewis's treatment of the problem of nuclear deterrence, Lewis nudges us to consider separately the questions of instrumental rationality and of morality, which is one of the views placing Lewis's ethics and social philosophy on the non-consequentialist stride. On this central view of Lewis's ethics and sociopolitical philosophy, I now analyze select essays amongst those mentioned in **subsection 7.1.**

1. In the essays "**Utilitarianism and Truthfulness**", "**Mill and Milquetoast**" and "**Illusory Innocence?**", Lewis argues for a non-consequentialist treatment of problems posed by utilitarianism.

- Using an example of coordination problems between two act-utilitarians, in the first essay Lewis recognizes the problem of disutility of expected utilitarianism. Lewis explains that the agents' abidance to truthfulness for coordination purposes is dependent upon truthfulness bringing about the optimally expected outcome for maximizing utility, which brings forth the problem of disutility. Lewis raises his concern that the agents will only abide to truthfulness for purposes of coordination, if they think that truthfulness will bring about the optimal outcome, that is maximizing utility. As Lewis (1986c: 341) argues:

"Knowledgeable and rational act-utilitarians would have no reason to expect one another to be truthful, not even when the combination of truthfulness would have good consequences; so they would forfeit the benefits of communication."

Moreover, Lewis (1986c: 341) stresses that it is most likely they would forfeit the benefits of all conventions, such as the conventions of truthfulness and of promise-keeping, which serve the purpose of coordinating our actions for the pursuit and fulfillment of common interests.

- In line with the above critique, in the second essay "Mill and Milquetoast", Lewis (2000 [1989b]: 163) reveals the flaws of Mill's neutralism as a constraint on utilitarianism when adopting toleration as a neutral principle of utilitarian reasoning. Using the hypothetical example of religious conflict, amongst religious factions, or between a religious faction and atheists, involving an inquisitor of utilitarian views and a group of heretics, Lewis shows that the utilitarian balance of costs and benefits in applying toleration will not succeed in such a case. Lewis (2000 [1989b]: 163-164) argues that, in Mill's neutralist rule-utilitarian approach, one wins this case by applying toleration, only if the benefits of toleration outweigh the costs. Thus, the inquisitor will be tolerant towards the heretics as long as toleration maximizes utility (Lewis, 2000 [1989b]: 167). In any other outcome, Mill's toleration of heresy will not be persuasive.

Admitting his lack of fundamental objections to broadly utilitarian reasoning (Lewis, 2000 [1989b]: 163-164), Lewis does not reject here a consequentialist interpretation, by arguing that the heretics' damnation is harm in precisely the perspective that Mill wanted to prevent: "it is the utter absence and the extreme opposite of human excellence and flourishing" (Lewis, 2000 [1989b]: 168). With reference to Mill's term of dead dogma, Lewis further makes the argument that Mill's account of deadness of dogma as a harm is justified only if received opinion is true, while admitting that it is worse to hold error as dead dogma, "rather than in a real and heartfelt and reasoned way" (Lewis, 2000 [1989b]: 164).⁹⁰

- The above positions are exemplified in "Illusory Innocence?", where Lewis considers the following problem of normative ethics. Lewis describes an encounter with a stranger, who had an accident, and urgently needs to be taken to hospital because of severe injury, or he risks losing one of his limbs. Taking the stranger to hospital will incur costs of time and a considerable amount of money, while it requires resources that do not belong to you. Lewis compares this hypothetical incident with the more common request from a charity for one's donation, to improve emergency medical assistance in a distant place, which would save people's lives. The charity contribution will

⁹⁰ Also in his essay on the choice of employee appointments in universities, "Academic Appointments: Why Ignore the Advantage of Being Right?", Lewis criticizes a utilitarian approach based on the candidates correctness of opinions for the advancement of knowledge, because it will lead to conflict between those who hold the correct and those who hold the false opinions.

cost less in money, time and bother, than saving the injured stranger. Why would most of us take the stranger to hospital, but not donate to charity?

In his interpretation, Lewis (2000 [1996c]: 153-154) here considers whether our commonsensical moral opinions are possibly right. Most of us would judge as a serious wrongdoing not coming to the occasional, or rare, aid of the injured stranger, while we see contributing regularly to a charity dedicated to medical assistance in a distant place as a recommendable act of optional generosity. Lewis examines an extreme utilitarian approach, like Unger's, which argues that it is a requirement of ordinary morality that we offer, not only our annual modest donation to potentially save many lives, but our inexhaustible resources. Siding with folk ethics, Lewis puts forth the question that, if it is seriously morally wrong to save the life of one distant person by donating to charity, why is it not seriously morally wrong to save the life of another one, and the next one, and the next, and so on. And what happens after you have contributed your fair share and the rest has been fairly divided amongst the world's most affluent? Lewis (2000 [1996c]: 154) shows that the argument only shuts off "when you have so little left that it becomes doubtful whether you can live to give again another day".

In my evaluation, in the aforementioned essays Lewis construes broadly his ethics and social philosophy as non-consequentialist, as we also saw in **subsection 7.2**. If we were to follow a consequentialist evaluation of Lewis's ethics in the plurality of worlds, we would come across the following claim: "Only if morality consists of maximizing the total of good, absolutely regardless of where and to whom the good may accrue, can it lose its point because the sum total of good throughout the plurality of worlds is non-contingently fixed and depends not at all on what we do" (Lewis, 1986a: 127-128). This claim does not accord with common sense, while Lewis defends it further in his anti-utilitarianist claim as consisting only in the subversion of a "truly universalistic ethics" (Lewis, 1986a: 128). For Lewis (1986a: 128), if, in the above recognition, what collapses is primarily a philosopher's invention of an ethics of our own world that is "already far too universalistic", there is not much harm done.

In the above mentioned moral philosophical examples, as we saw in **subsection 7.2**, it is clear that Lewis (2000 [1978]: 34-35) argues against certain strands of utilitarianism, such as what he calls "radical utilitarianism", as well as "anachronistic utilitarianism" (Lewis, 2000 [1989b]: 172) and its distant relative "the

'saint' utilitarian" (Lewis, 2000 [1984b]: 206), because they are remote from common moral sense. Whether this view offers an explanation of Lewis's critique of utilitarianism is, however, not sufficiently clear. As we saw at the beginning of this subsection, Lewis also encountered the problem of differentiating between moral common sense and moral problems in canonical ethics regarding folk moral psychology, which he sees as a combination of psychology and ethics. Lewis (1986a: 128) argues that his version of modal realist ethics is not contrary to common sense: "modal realism creates only a problem for utilitarians of an especially pure sort", as Lewis (1986a: 127) explains.

I evaluate this Lewisian main argument to be also a metaethical argument about modality in folk ethics. It is a matter of whether the opinions about how to maximize utility are wrong or right, not whether it is wrong or right to aim for maximizing utility. This Lewisian view is, for instance, evident in "Mill and Millquetoast" (Lewis, 2000 [1989b]: 169). Again, in his critique of act-utilitarianism, Lewis raises a problem of modal thinking, which he identifies with utilitarian approaches that do not succeed in their intention of maximizing expected utility, rather than questioning the moral principle itself of utilitarianism as a consequentialist moral philosophy.

2. Challenging further consequentialism, which, as we saw, is a strand of utilitarianism, the moral philosophical question of deterrence appears in Lewis's essays on questions of applied ethics, such as the problems of nuclear deterrence and deterrence as a judicial function via the penal system. Lewis examines the political philosophical problem of hypothetical nuclear deterrence in the essays "Devil's Bargains and the Real World" and "Buy Like a MADman, use like a NUT", to argue in both essays against deterrence as an irrational response to the threat of nuclear attack by inflicting pointless harm. In addition, Lewis evaluates practices of deterrence and punishment in the current penal system in the following three essays: "The Punishment that Leaves Something to Chance", "The Trap's Dilemma"⁹¹, and "Do We Believe in Penal Substitution?". Considering Lewis's views on utilitarianism and consequentialism, as discussed previously in this subsection, I now analyze two of the aforementioned essays on the question of deterrence, one on each specific problem.

⁹¹ In "The Trap's Dilemma", Lewis uses deontic logic to argue that a policeman, who has sworn a professional oath to obtain a conviction, will be necessarily dishonest, whether he breaks the oath or not, if the only way to obtain a conviction would be to lie.

- To start with, in the essay "**Devil's Bargains and the Real World**", Lewis (2000 [1984b]: 201) presents us with a definition of deterrence as the dissuasion of someone from doing harm, by forming a genuine, effective and conditional intention of retaliation, if he or she does harm. However, Lewis (2000 [1984b]: 202) challenges consequentialist presumptions in his argument that in some cases, "it is in truth right to form an intention that it would be wrong to fulfill"; for instance, if retaliation would serve no other purpose but inflicting further pointless harm (Lewis, 2000 [1984b]: 201). In support of his non-consequentialism, Lewis examines Gauthier's position, which he is agreeable with, apart from the crucial points where Gauthier appeals to rationalizing or moralizing deterring behavior. Lewis (2000 [1984b]: 204) claims instead that, in certain extreme cases of deterrence, it may be "rational to commit oneself to irrational behavior". Lewis makes this statement in view of the hypothetical example where military personnel has to obey orders to launch nuclear strikes in defense of their country (Lewis, 2000 [1984b]: 205-206), but not just for the sake of vengeance, because such a motivated action would inflict off-target harm (Lewis, 2000 [1984b]: 207). This simply means that Lewis (2000 [1984b]: 203-204) considers separately the question of the rationality (or morality) of intention and the question of the rationality (or morality) of action, on their own merits – rather than, for instance, reasoning causally from the first to the second, as Gauthier does, or vice versa.

It is important to clarify here that Gauthier's argument (1984: 477) implies the intention of maximizing rationality, which Gauthier presumes has a cause and effect relationship with acting upon the intention; as Gauthier (1984: 480) argues, "if it is rational to form the intention, it is rational to act on the intention". Equally, Gauthier (1984: 483) argues that, "it is rational to execute an intention if and only if it is utility maximizing to form it". In the specific problem of nuclear deterrence, taking the extreme hypothetical example of a potential nuclear catastrophe, Lewis (2000 [1984b]: 204) explains that he accepts the policies of nuclear deterrence as right, but rejects "the actions they (conditionally) require as wrong", in this way rendering his argument consistent, because it is about two different things. Conversely, Gauthier (1984: 480) defends the rationality of nuclear policies, including nuclear retaliation, by adding the clarification that "among rational and informed actors, a policy of pure and simple deterrence is not rational, although it

may be rational as part of a larger policy directed, among other things, at the obsolescence of deterrence", which means that rational actors would be expected to move "not unilaterally, but mutually, beyond deterrence".

As Lewis (2000 [1984b]: 203) points out, Gauthier and himself are largely in agreement on the broader argument of nuclear deterrence, based on the maxim: "To seek Peace, and follow it' and 'by all means we can to defend ourselves'" (Gauthier, 1984: 495). As we saw, their disagreement rests on their characterization and justification of the reasoning behind practices of deterrence, based on their differences in the use of the terms rationality and irrationality. Gauthier (1984: 486-487) further examines an objector's view on the connection between utility maximization and rationality, although, as he admits, "the two preferences are logically and actually quite distinct". In Gauthier's view (1984: 487), such a connection between conditional intentions and their implementation in cases of deterrence is not valid, "because adoption of the intention affects one's expected utilities by affecting the probability that the condition for implementation will be realized". Furthermore, since Gauthier (1984: 487) takes that an irrational action is an action that "should not be, or have been, performed", if we accept deterrence policies, then we cannot claim that the actions required for their implementation simply cannot be performed.

One could argue here that a position upholding as rational a policy, which requires impossible irrational actions, results in a sort of moral fictionalism, which I discussed previously in **subsection 7.2** and which Lewis himself rejects. Gauthier's suggestion (1984: 488) is that the agent who assesses their actions in the context of aforementioned policies might be better off adjusting their performance for maximizing the probability-weighted sum of their utilities, rather than non-performing. Lewis (2000 [1984b]: 210) sides with a rather agnostic view on what ought to be done on the practical level of public policy regarding apparently paradoxical cases, such as nuclear deterrence, which are interesting on the theoretical level, however disconnected from the actual practices of say nuclear deterrence. Pointing at the disutility of philosophical discourse on public affairs, Lewis (2000 [1984b]: 211) defends nuclear deterrence policies as non-paradoxical by applying his modal realism on our common sense moral opinions in the extreme case of a nuclear war (Lewis, 2000 [1984b]: 211-218).

- Secondly, in the essay "**The Punishment that Leaves Something to Chance**", Lewis (2000 [1987]: 241-242) applies probabilistic reasoning to conclude that his proposed practice of perfect crime reenactments is as close as it can be to present practices, which, in comparison, are more just than chancy punishment by the "penal lottery". Lewis puts forth the following hypothetical problem drawn from the penal system. Two agents act out of malice to cause harm, without justification or excuse. One has killed; he is considered guilty of murder and receives the death penalty. The other one has not killed; he is considered only guilty of attempted murder and receives a short prison sentence. Lewis asks the question why, in our penal system, we give a more severe punishment to an agent, who has succeeded in his attempt to cause harm, rather than to the one who has put a victim, knowingly and wrongfully, at risk of harm, but has failed only by chance to cause any actual harm.

Lewis examines different common rationales about deterrence rather than punishment, which may also defend the above described practice as just, presuming our agreement that both agents have been equally wicked in their desires and that they have engaged in conduct, which should be prevented by deterrence. Firstly, Lewis rejects theodicy and prudence as sufficient and just reasons for sustaining our penal practice. He also rejects positions that are either concerned with punishment as minimum deterrent, or with the deterrence of second attempts, because they are unlikely to be successful at deterring all agents and they cannot guarantee the degree to which deterrence will be successful. Finally, he examines rationales regarding the "moral luck" of the agents and the whole-heartedness, or half-heartedness, of their attempts to cause harm, in order to also reject them based on consequentialist arguments. Lewis (2000 [1987]: 241-242) proposes the "penal lottery" by enactment, which is much like the present practice, instead of what he calls simply the "penal lottery", that is punishment by chance: either by reenactment or by drawing straws.⁹² In this case, we use the original crime at its perfect reenactment. If the victim dies, the criminal dies too, but if the victim lives, the criminal gets a short prison sentence. Only, the lottery must not settle "so soon that the criminal may know its outcome before he [she] decides to commit his [her] crime" (Lewis, 2000 [1987]: 242). The practice is not necessarily just, but it is not necessarily unjust either.

⁹² Nolan (2005: 191) describes the problem of the "penal lottery" in lay terms, as letting "[...] the criminal get off scot-free just because their friend unjustly took the rap".

In my evaluation, in his treatment of problems of normative ethics with a social and political focus, such as the tacit treaty of toleration and the application of deterrence in the penal system, Lewis formulates a consistent critique of utilitarianism, and hence of consequentialist moral philosophical theories. Lewis's method consists in problematizing utilitarianist approaches to hypothetical cases, by restating the cases to offer counterarguments and counterexamples, yet without disputing the normative ethical principles themselves. Notably, about Lewis's modal realist account of utilitarianism, objectors have stated "the concern that his modal realism would produce moral indifference" on the utilitarian premise that evils will be actualized in some other possible world anyway, if they don't occur in ours (Nolan, 2005: 189). We can trace Lewis's responses to similar arguments by McMichael and Unger with reference to his use of hypothetical examples, which show that Lewis counter argues such objections by appeal to: i) proposals for application of value structures and deontic principles at worlds, but not across all worlds and, or, on utilitarian principles based on the neutralizing principle; ii) common sense morality, which objects to doing good regardless of who is the beneficiary. I continue with these arguments in the following **section 8**.

8

Challenging the canon: Lewis's positioning in respect to canonical accounts

As seen so far, Lewis's contributions to ethics and social philosophy challenge, to a certain extent, canonical accounts in ethics and metaethics, as well as in social and political philosophy. This preliminary assessment has come out of my analysis of Lewis's ethics and social philosophy in respect to canonical accounts, which he addresses primarily in his essays, but also, secondarily, in his seminal works on convention, counterfactual conditionals, and his metaphysical thesis of the plurality of worlds. I now concentrate on a more in-depth evaluation of Lewis's contributions to relevant select canonical discourses, namely: determinism and non-determinism; compatibilism and incompatibilism; realism, quasi-realism and modal fictionalism; conventionalism; and utilitarianism, which is a strand of consequentialism. Besides offering a critical positioning of Lewis's ethics and social philosophy in respect to canonical accounts, in this section I argue that Lewisian modal realism is indispensable to the above, by placing emphasis on Lewis's modal realist responses to canonical views in my consideration of other modal realist counterarguments.

I. Lewis's treatment of problems of **determinism** and **non-determinism** in view of the Lewisian modal realist account of possible worlds is evident in Lewis's addressing of questions of causal relations and nomological causality: for example, as seen in **section 4**, regarding Scriven's problem of predictor and avoider, as well as Nozick's interpretation of Newcomb's problem. Lewis (1973b: 559) explains that, by determinism, he does not mean "any thesis of universal causation, or "universal predictability-in-principle", but rather gives the following definition:

"[...] the prevailing laws of nature are such that there do not exist any two possible worlds which are exactly alike up to some time, which differ thereafter, and in which those laws are never violated" (Lewis, 1973b: 559).

Lewis (1973b: 559) also highlights the view that a counterfactual analysis holds an advantage over a regularity analysis, because "it allows undetermined events to be caused", in line with Lewis's modal realist view.

To understand better Lewis's position regarding the question of determinism and non-determinism, we can examine his take on another modal realist interpretation of possible worlds, such as van Inwagen's. We analyze the Lewisian principles of modal realism (1986a: 86) in possible worlds as follows:

- (1) "Absolutely every way that a world could possibly be is a way that some world is."
- (2) "Absolutely every way that a part of a world could possibly be is a way that some part of some world is."

We encounter the valid objections that both principles are contentless, because they only point at:

- (1) "Every world is identical to some world."
- (2) "Every part of a world is identical to some part of a world."

For van Inwagen, these principles do not say anything about abundance or completeness in worlds, yet they could be true regardless of the number of worlds, which could also be none (Lewis, 1986a: 86). Lewis (1986a: 86) recognizes that in his own version of modal realism, it is advantageous, in virtue of the economy of the theory, "to identify ways a world could possibly be with worlds themselves". For Lewis (1986a: 86), this means that a world and the maximally specific 'way' that world is are identical. This identification allows for abundance and logical completeness. Additionally, 'a possible way' a world is, is a non-empty set (Lewis, 1986a: 87); something that, responding to van Inwagen, Lewis demonstrates by substituting the maximally specific 'way' a world is, which constitutes a Lewisian modal realist world, by a unit set. Lewis (1986a: 87) recognizes that his principles are not those of plenitude, however, in his modal realist view, this is only a trivial concern, as long as there are enough possibilities and no logical gaps.

Let us now consider that, for 'ersatz' modal realists, as Berto and Plebani (2015: 173) argue, representation by worlds is a maximal state of affairs of actuality (Plantinga, Van Inwagen). As Lewis (1973b: 81) argues, maximal properties by world-

selection functions based on the notion of comparative similarity by means of accessibility relation (Stalnaker), introducing ways things might be as something like maximal structured properties of actuality as a whole. For so-called 'ersatz' modal realists, the actualized world is the unique state of affairs, or world property, which is instantiated. This means that of all worlds, only one obtains, which is the actualized world. Nevertheless, we understand that in Lewisian modal realism, representation by worlds is the maximal specific 'ways' things might be (Lewis, 1973b: 84), as I argued in **section 2**.

Van Inwagen argues that Lewisian worlds encounter a problem in their definition by necessarily equating a world with the maximally possible specific 'way' that world is, which does not preclude a world from not existing. Lewis's response by substitution of a maximally specific 'way' a world is by a unit set, which demonstrates that a Lewisian possible world is a non-empty set despite being a trivial principle of plenitude, counter argues van Inwagen's problem of the non-existence of possible worlds. As with Plantinga, also for van Inwagen, representation by worlds is a maximal state of affairs of actuality: only one world is actualized by instantiation, hence there is the problem of the existence of possible worlds in abstract terms. On the other hand, Lewisian worlds (1986a: 103) are alternative possibilities, rather than parts of one big disunified actuality. For Lewis (1986a: 129), the main difference between ours and other possible worlds is that our world is the world we inhabit. Yet this Lewisian notion of inhabitation at a world has nothing to do with the notion of actualization of a world in 'ersatzist' terms.

We saw that Lewis (1986a: 81) accepts that possible worlds are the same in kind as ours, while he also accepts that, in his modal realist terms, the laws of physics – but not the mathematics – may differ at different possible worlds (Lewis, 1986a: 108-109). This view is consistent with Lewis's view of physics as a science dealing with physical properties (1986a: 60); as he claims, "[...] in other worlds, where physics is different, there will be instances of different fundamental physical properties, alien to this world". In fact, Lewis (1973b: 90-91) argues that mathematical 'ersatz' constructions of possible worlds are too dependent on current knowledge of physics, although they are not concerned with physics proper, but rather with metaphysics done by physicists. Lewis (1986a: 103) locates the difference between seeing possible worlds "as genuine alternative possibilities", rather than "as parts of one big disunified actuality", on this dependence of

'ersatzism' on current physics, which has an effect on Lewis's account of causation. In this view of possible worlds, Lewis (1973b: 558) understands causation among events in the everyday sense, by defining causation as instantiation of regularities following a counterfactual analysis (Lewis, 1973b: 557), in which Lewis (1973a: 563) distinguishes between counterfactual and causal dependence from what he terms as 'nomic' dependence: for instance, in Scriven's and Nozick's problems.⁹³ Lewis examines 'nomic' dependence in the case of raising one's arm and the causation of a law-breaking event, for which a 'nomic' deterministic explanation would be that any 'divergence' requires some violation of the actual laws of nature (Lewis, 1973a: 566-567). Instead, as shown, Lewis (1986a: 103) puts forth the argument, which can be taken as a refinement of the 'nomic' deterministic explanation of 'divergence', that some natural, non-ad hoc break in possible space-times has occurred, which adheres with the mathematical universe.⁹⁴

2. Furthermore, taking on the question of **compatibilism** and **incompatibilism**, as seen in **section 7**, several Lewisian moral problems put forth the question of determinism in respect to free will. It is worth examining a modal realist formulation of the problem of incompatibilism, by means of a 'necessity' and

⁹³ This Lewisian view is consistent with Nolan's interpretation of Lewis's "laws of nature" (2005: 86-87) as contingent only upon what is discovered by some branches of physics, while what is discovered by biology, chemistry, biology, linguistics, or archaeology are certainly not "laws", but just "nomically contingent facts" about our world, which accords with Lewis's view of the contingency of empirically investigated scientific knowledge (Lewis, 1973b: 91). In this sense, Nolan (2005: 9) argues that Lewis seems "reductionist", as far as he supports the view that "[...] the basic laws of nature can be cast in terms of the patterns of fundamental physical properties and relations". Psillos (2006: 155) also comments on the Lewisian view that, although, for Lewis, the basic thesis of the deductive nomological model has not been discredited, it has been "wounded". This means that, in Lewis's view, a deductive-nomological explanation of an individual event's causal history, although "both possible and most complete", this is a more like a "chimerical" ideal (Psillos, 2006: 156).

⁹⁴ Lewis (1973b: 105-106) also argues that past and future systems are more like deontic systems, neither centered, nor weakly centered, while they do not fulfill the conditions of absoluteness and universality. As discussed in **subsection 6.1**, footnote 75, Lewis refutes the Stalnakerian Limit Assumption. For Lewis (1973b: 105), the view of time's linearity does not guarantee the Stalnakerian Assumption, but rather makes Stalnaker's Assumption equivalent to the Limit Assumption. In Stalnaker's account (1990: 324), the view of time's linearity influences his notion of a "branching time structure", defined as "paths through a branching tree of temporal nodes", in which possible worlds are conceived as "partly overlapping paths through the tree"; for more on this view of time, see (Gamut, 1991: 40-44).

Based on the theory of quantum mechanics, Lewis's own view does not reject the branching time structure (Lewis, 1986b [1973]), but redefines it as quantum-mechanical branching (Lewis, 2004b), in this way putting forth the notion of 'divergence' in possible worlds, by taking on the relation of the comparative similarity of worlds, which enables non-linear accounts of time and causality (Lewis, 1979b). In this view, Lewis's support of 'genuine' four-dimensionalism (2002 [2004]: 9) could be interpreted as an explicative concept for his earlier view of possible worlds as "us and all our surroundings" (Lewis, 1986a: 185). Kripke (1980: 45) also compares Lewis's counterpart theory with the Wheeler-Everett many-worlds interpretation of quantum mechanics (MWI).

a 'conditional' rule following van Inwagen's modal realism (2008: 452-453). Drawing from the basis of van Inwagen's own formulation of the same problem (2008: 455-456), which is the notion of an "untouchable truth", such as say a law of nature, playing the role of every true proposition, it is not difficult to see why his questioning of the incompatibilist argument must be settled on whether the conditional rule is valid. For example, reformulating the problem of an agent's free will regarding a law-breaking event in van Inwagen's modal realist terms, the proposition would hold as: "I am never able to do anything else than not breaking a law other than not breaking a law"; which is a consequentialist argument. The Lewisian modal realist proposition reformulated following van Inwagen would be:

- (1) "I am not able to break a law." ('necessity' rule is valid)
- (2) "If I had (or had not) raised my arm, I would not have been able to break a law." ('conditional' rule is invalid)

In Lewis's modal realist terms, I am not able to break a law, whether I raise my arm or don't raise my arm; yet a law-breaking event occurred. We can now consider Lewis's response to the consequentialist argument, whether the word "untouchable truth" is vague (Van Inwagen, 2008: 456). Since Lewis starts from a different premise, by questioning the notion of the "untouchable truth" in reference to a law of nature, he can defend his view of the occurrence of a so-called "divergence miracle" (Lewis, 1986a: 130), which is neither determined, nor caused by the agent's free will. Lewis upholds a compatibilist and non-determinist position, as well as counter argues the consequentialist argument, in response to the free will and determinism incompatibilist argument, by providing a counterexample to van Inwagen's preliminary validation of the 'conditional' rule.

To understand this seemingly paradoxical Lewisian position, we need to bear in mind Lewis's own use of the theoretical terms compatibilism and incompatibilism, concerning: i) Lewis's view of the 'compatibility premise' used in his response to Scriven taking an indispensable set of ideal conditions for both predictor and avoider, which involves them having "perfect knowledge of data and laws, must be able to do predictive calculation with perfect reliability, and must also be able to do enough calculation" (Lewis, 2000 [1966]: 246). Lewis (2000 [1966]: 250) argues instead that, in fact, we only deal with an approximation of this set of ideal conditions; ii) Lewis's view of compatibilism and incompatibilism in respect to one's character mentioned in the problem of freedom and free-will theodicy.

Lewis (2000 [1993]: 109) argues that, in view of compatibilism, our choices are free insofar as they manifest our characters, including our beliefs and desires, and bypass causal chains beyond our characters. So we are free, but to a certain degree predetermined by our characters. Lewis's view of incompatibilism (2000 [1993]: 109) is that our choices are free, only if they have no determining choices outside our characters; not even causes that determine our choices through our characters. These Lewisian views further explain Lewis's non-consequentialist argument in the example of the law-breaking event regarding one's free will and nomological causality. However, conversely to his view of the contingency of empirically investigated scientific knowledge (Lewis, 1973b: 91), in his value theory, Lewis (2005: 230) recognized the need for empirical information "about what we are in fact disposed to value under the appropriate circumstances".

3. Continuing with questions of **realism**, I consider here Plantinga's modal realist argument (2003: 212-213) that Lewis is a modal anti-realist, or at least "a realist of an interesting kind", but not a modal realist, because of his modal reductionist approach to possible worlds and individuation,⁹⁵ as described below:

"On his theory, as I see it, there are no propositions, states of affairs, possible worlds, essences or objects with essential and accidental properties; what there are instead are concrete objects and set theoretical constructions on them, some of which play roles formally similar to the roles in fact played by the phenomena of modality if the modal realist is right." (Plantinga, 2003: 213)

Although Plantinga (2003: 212) recognizes that Lewis is a realist in terms of "a plurality of maximal objects", he sees the Lewisian position, which models them and their properties "in concreta and sets" (Plantinga, 2003: 203), as an alternative to Quine's rejection of the disorderly crew of *possibilia* in subtle modal reductionism.

There are obvious differences in the modes of representation between the Plantingan modal realist and the Lewisian 'genuine' modal realist interpretations. First, Plantinga is mistaken in his view of the Lewisian *de re* representation and counterpart theory. For Plantinga, a property (i.e. Socrateity, or being Socrates) is

⁹⁵ As Nolan (2005: 9) mentions, Lewis describes himself as a "reductive materialist", but he is certainly not a reductionist, in the sense that mental processes are "reduced" to physiological processes, or that biological systems are "reduced" to physical systems, so we can claim that there are no mental processes or biological systems. An example of Lewis's dismissal of this sort of "reductionism" is his view on the psychophysical supervenience thesis; see **Appendix I**, "Objections and Replies", I.

not the property of being identical with the person, who in the actual world W, is Socrates; as Plantinga (2003: 83) argues,

"[...] it is not the property of being that person. It is the property that could have been had by someone else; roughly, a property that is unique to Socrates and his counterparts."

However, individuation in Plantinga's terms is not individuation in Lewisian *de re* representational terms. As explained, in Lewisian modal realism, individuals are wholly part of one or more possible worlds. In the example of Socrates, having a property, such as Socrateity, means that Socrates has that property at that world W, of which he is wholly part, and by means of having that property he is wholly part of the world W, where he is Socrates and has Socrateity. The same applies to Socrates' counterparts at W_n worlds. In my view, this difference in modal representation between Plantinga and Lewis does not account for modal reductionism on Lewis's part. Conversely, one can argue that Plantinga's own more generalized view of individuation is reductionist, in terms of ascribing a property common to Socrates and his counterparts to anyone at any world, instead of taking individuals as world-bound, which means, as I explained in **subsection 5.2**, that properties of individuals come with their surroundings.

Furthermore, taking Lewisian *de se* representation with Lewis's dispositional theory of value (2000 [1989a]: 89), one can support Lewis's "irrealist" view with regard to the ontological necessity of values:

"If we amend the dispositional theory by inserting 'necessarily', we can be much more confident that the 'values' it defines would fully deserve the name – if there were any of them. But it is hard to see how there could possibly be. If a value, strictly speaking, must be something we are necessarily disposed to value, and if our dispositions to value are in fact contingent, then, strictly speaking, there are no values." (Lewis, 2000 [1989a]: 90)

We saw that Lewis (2000 [1989a]: 93) argues that values exist "loosely speaking". In *de se* modality, the belief-desire concerns an individual's own place at a world, therefore value is contingent with respect to that world. Nonetheless, Lewis does not support error theory. In reference to Mackie, Lewis (2000b: 90) argues that we should not amend the dispositional theory of value to accept an extreme contingency of value, in the sense that we require values "to be all we might wish them to be". Lewis (2005: 320) explains that his realist position in ethics "is a kind of

subjectivism, though remote from simpler kinds of subjectivism" and that it is "conditionally relative", based on differences to underlying value dispositions, which indicate speakers in a population in terms of 'we'.⁹⁶

Turning to Unger's problem concerning an apparent dilemma of benevolence and charity, which I analyzed in **section 7**, we can apply modal realist interpretations. In Lewisian terms, *de re* representation requires that the given stranger has the property of being wounded at world W, which is our world. In Plantingan terms, being wounded is a property that anyone can have at any world.⁹⁷ Common sense morality requires that we help the stranger, because we are grouped together with him in a particular situation at a particular world W, which is also our world (Lewis, 2000 [1996c]: 155-156). We are less inclined to help infinitely any wounded stranger at any possible world, which is a common sense moral commitment.⁹⁸ Hence, one can argue that a Plantingan interpretation is removed from common sense. Evidently, this moral problem also poses questions of representation of different possibilities, which Lewis outlines in view of his critique of modal realists; for instance, those he characterizes as 'magical ersatzists', including van Inwagen and Plantinga, as seen in **section 1**.⁹⁹

4. Further on the philosophical question of realism, we saw that Lewis counter argues problems concerning **modal fictionalism** and **quasi-realism**. I draw upon the Lewisian view on modal fictionalism (1978; 2005) to review one of the problems analyzed in **subsection 7.2**, in the context of Nolan's critical analysis (1997) of different modal fictionalist positions. Lewis (1986a: 141) certainly does not support a modal fictionalist view of worlds:

"Our own world cannot be such a world; for if it is really a fiction we are dealing with, then the act of storytelling at our world was not what it purported to

⁹⁶ Nolan's argument (2005: 180) against the description of Lewis's position as purely "subjectivist" also accords with Lewis's rejection of the error theory. As Nolan (2005: 180) argues, although Lewis thinks that values are settled by our dispositions, as I discussed in **section 4**, he does not support the view that "whatever anyone's opinion is about what is valuable, that is automatically correct".

⁹⁷ This view is also consistent with the actualist position, which, for Lewis (1986a: 97), applies to everything.

⁹⁸ Quinn (2004: 712) also comments on Lewis's critique of Unger's "shocking conclusion that, as long as there is an inexhaustible supply of distant strangers whom we could rescue from urgent need, ordinary morality requires us to give almost everything we possess or can acquire to assist them". For Quinn (2004: 713), in this case Lewis does not offer an argument against utilitarianism, but he rather supports "basic values to which we are already committed", which Unger attempts to challenge. Unger (2004) has explored problems of representation, in regards to what he terms as 'the experiential problem of the many' and the problem of 'too many real choosers', in the context of Cartesian metaphysics.

⁹⁹ For a more detailed account of the differences between Lewis's modal realism and other modal realist accounts, which Lewis terms as 'ersatzist', see **Appendix I**, "Objections and Replies", 3.

be. It does not matter if, unbeknownst to the author, our world is one where his plot is enacted." (Lewis, 1983f [1978]: 266)

Besides the Lewisian metaphysical conception of possible worlds as a substantive thesis of the plurality of worlds, the Lewisian modal realist position is not fictionalist, because of the Lewisian philosophico-logical approach to modal realism (1986a: 115), which, as I discussed in **section 1**, follows a combined method of informal reasoning and reasoning from mathematical axioms. The 'genuine' modal realist comes across several problems within modal 'ersatzism', such as problems of representation and of the existence of abstract and concrete entities, but also epistemic problems of verification and truth value.

In relation to counterarguments by modal fictionalism, which methodologically adopts the employment of "the notion of fiction to explain some of our talk about modality, and particularly our talk about possible worlds", as Nolan (1997: 259) has argued, modal realism can easily counter argue inherent problems in modal fictionalism, such as the elimination of modality (Nolan, 1997: 262), but also the artificiality infecting modality (Lewis, 1983f [1978]: 266; Nolan, 1997: 264) and the ontological inadequacies, which are present in strong fictionalism (Lewis, 1983f [1978]: 274; Nolan, 1997: 265). Significantly, as we saw in **subsection 7.2**, Lewis distinguishes between the modal reasoning of our common sense modal opinions, which he associates with folk moral psychology, and the false argument that "reasoning about truth in fiction is very like counterfactual reasoning" (Lewis, 1983f [1978]: 269). Lewis exemplifies this position in his analysis of the function of folk fiction, which consists in fictions told as known fact in worlds, yet they originate in the collective belief worlds of the community of origin. On this basis, Lewis (1983f [1978]: 273) argues that considering any overlap between these two sets of worlds is a mistaken view.

On the quasi-realist position, we saw that Lewis puts forth the counterargument of moral fictionalism, in which morality is retained with identifiable values but treated as fiction (Lewis, 2000 [1989a]: 93). I explained that Blackburn's response to this argument is that quasi-realism must not be confused with moral fictionalism, that is with morality as make-believing (Blackburn, 2005: 327). In his reply to Sturgeon's analysis of moral fictionalism, Blackburn (1991) offers additional explanations to his quasi-realist position against moral fictionalism, by showing the difference between moral fictionalism, convention and ethics.

Blackburn (1991: 3) argues that a projective theory of moral psychology has two sides: an 'input' side, which comprises of the "natural features of the world to which the subject is responsive to forming an ethical commitment", and an 'output' side, which is "the kind of commitment formed". Significantly for projectivism, these need to be identified, "initially at least", without our reliance on a moral vocabulary (Blackburn, 1991: 3). This means that, when the projectivist eventually introduces a moral principle or term as a given 'input', his moral explanations do not lose their ground on 'realism' (Blackburn, 1991: 3). Blackburn's definition bears implications for the range of outputs we encounter in examining individual members of a community; when, for instance, we hold constant as an 'input' the belief that a society is governed by a Rawlsian principle (Blackburn, 1991: 4). Here Blackburn emphasizes that, while meaning, say of a moral term, can be 'conventional', that is conferred in the term by an authority that we institutionalize (Blackburn, 1991: 5), ethics does not rely upon a convention as such, because there is no single authority when we are treating "essentially contested concepts" (Blackburn, 1991: 6). As Blackburn (1991: 8) notes:

"Meaning is only properly talked of where we have *convention*. But there is no convention governing the selection of standards in ethics: someone who approves of the wrong things is not unconventional in the way of someone who uses a word wrongly. A group cannot put pressure on an isolated moral view by claiming that it is a convention that other things are good or bad. Nor is there a convention governing the actual output: someone who feels a different way about the importance of something (that may be described as justice or virtue) cannot be brought into line by citing meaning."

In this view, Blackburn (1991: 7-8) questions dispositional theories of value, but also efforts to universalize moral values derived from the meaning of moral terms.¹⁰⁰ Blackburn's account (2006: 245) is consistent with his quasi-realist distinction between ethical or normative facts, which can seem "metaphysically strange and epistemologically demanding", and our moral attitudes or practical stances, which Blackburn (2006: 245) equates with the mental states "we voice as we communicate

¹⁰⁰ One can argue that philosophical debates relevant to ethics and moral philosophy on the role of the meaning of terms derive from the classical tradition's critique, including Frege and Russell, of Mill's argument for denotation versus connotation in the use of 'singular' names, 'general' names and 'proper' names (Kripke, 1980: 26-27; 127-128; 134-135). Lewis (2000 [1992]: 151) has placed emphasis on the significance of debates around meaning, in his response to Hawthorne on the problem of actual and possible languages in a population by reference to Kripke and Wittgenstein.

values or moralize to each other", which are dependent, amongst other things, on language usage. In this scope, Blackburn's proposal (1991: 16) consists in dealing with ethics without helping ourselves to moral properties or explanations from the beginning, but "to see them as constructions, or [...] projections, regarding ourselves in the first instance as devices sensitive only to natural facts, and producing only explicable reactions to them". In adopting such a position regarding the ontology of values, for Blackburn (1991: 17), we do not understand moral properties, such as justice or goodness, and "the properties of just and good institutions or actions", as being opposed to each other.

In response to Blackburn's quasi-realist view, in his value pluralism at possible worlds, as well as in his separate treatments of problems of normative ethics, Lewis makes suggestions for virtue ethics. Lewis (1986a: 127) draws our attention to virtues included in acts of judgment, preference, conduct and attention, as well as in prudential thinking (Lewis, 1986a: 219). For Lewis (1986a: 127), these acts stand for 'egocentric' moral values, which a virtuous person, who presumably wants to avoid evil doing by actualizing evil, will want to consider. Lewis (1986a: 125) also shows that we can avoid evil doing because of indifference, by recognizing virtues, such as honor, respect, esteem, loyalty, and affection, in *de se* modality, which is the 'egocentric' want that "I myself should have a certain property".¹⁰¹ This Lewisian suggestion complies, as we saw in **subsection 7.2**, with Lewis's view that deontic ethics can be construed on any principles – for instance 'radical utilitarianism' – which can produce results far removed from and unacceptable by common moral sense. However, as Lewis (1986a: 125) mentions, attitudes *de dicto* and *de se* may collapse into each other:

"Sometimes, what one desires is that the world should be a certain way: that it should realize one of a certain class of possibilities for the whole world [...] To desire that the world realize some possibility within the class is to desire that the proposition be true. Call this 'desire *de dicto*'." (Lewis, 2000 [1989a]: 73)

I argue that the above Lewisian view on virtue ethics does not contradict Lewis's position on the non-essentialism of properties – aside from say the property of being human. Lewis supports a non-essentialistic view of properties in his value pluralism and in respect to his modal metaphysics, where properties are successfully or unsuccessfully ascribed to world-bound individuals by reference

¹⁰¹ For more on Lewis's theory of individuation, see **Appendix I**, "Objections and Replies", 1.

and according to a world; yet there are no necessary values *per se*.

Lewis here presents us with the problem of identity, which is relevant to ethics, deontological and virtue based alike. To highlight the significance of *de se* attitudes in his ethics, Lewis (2005: 320-321) in fact argues that his moral theory is not so much a theory of the 'right', but of the 'good' of values:

"Let us augment it as follows. When we are disposed to value a certain property, it may happen that having that property (or having it to the fullest possible extent) requires unconditional compliance with certain constraints on conduct [...] If so, I shall say that the constraints built into the value are the obligations associated with that value".

Thus schematically outlined, Lewis's virtue ethics offers an alternative interpretation, including the modals *de se* and *de dicto*, to canonical accounts of virtue ethics, also in contrast to deontic ethics. In this scope, Lewis presents a method, different to Blackburn's, of arguing about the ontology of moral values; however, in my view, neither of them diverts substantially from a basic realist position.

5. As seen in **subsection 7.2**, Lewis's contributions to **conventionalism** derive from his theory of convention, where he adapts his Humean interpretation of conventions to game theoretical thinking and an initial conception of his possible worlds modal realism. We saw in **section 5** that this move is an early indication of Lewis's support of value relativism with regard to worlds, as well as of his conventionalist approach to the notion of analyticity, which is relative to a population of language users and determines the possible language that is the actual language in a population. We also saw in **section 7** Lewis's clarifications in response to Jamieson regarding his own conventionalist approach to value theory, which highlight the significance of the role of salience in selecting conventions in populations. As we saw in **section 5**, Lewis (2000 [1976b]: 142) argues that his own concept of 'common knowledge' in a convention, which is relevant to a population, simply requires "knowledge of conditional preferences", but not of "conditionals about preferences". This view complies with Lewis's conditionally relativist view of value, which as we saw in **section 4**, is relevant to a population and established by approximation to canonical methods for the definition of value.

Arguably, Lewis's original contribution to conventionalism rests on his application of game theory as a methodological invention, which allows him to

carry out the task of social philosophy, despite Lewis's (1969: 3-4) lack of an explicit articulation of such an argument. Lewis's obvious innovation in the use of the concept of the 'common knowledge' of rationality in game theoretical terms (Lewis, 1969: 70) leads to his proposal that conventions need to abide to specific conditions, such as regularity in behavior, but also systems of mutual expectations and mutual preferences, which are shared as 'common knowledge' in a certain population (Lewis, 1969: 58; 78). In Lewis's conventionalist account (1969: 75), conventions are adaptable and substitutable (Lewis, 1969: 85), as far as the 'common knowledge' requirement is satisfied; as Lewis (1969: 57) argues, "a basis for common knowledge generates higher-order expectations with the aid of pre-existing higher-order expectations of rationality".¹⁰²

It is important here to note that Lewis (2005: 316) treats 'rationality' as an "elastic notion", supporting the view, as we saw in **subsection 7.2**, that it is only a trivial truth that moral beliefs automatically motivate the rational: "If someone sees fit to classify the pursuit of the right and good as an aspect of 'rationality', I am not as sure as I would like to be that he has exceeded his linguistic rights; and under that usage it is trivially true that moral beliefs automatically motivate the rational". The Lewisian take on the notion of 'rationality' also explains his view of conventionalism, which in Lewisian terms can also be defined as a "semantical game",¹⁰³ relying on a game theoretical treatment of truth (Hintikka, 1986: 236-237). This is evident in Lewis's study of social behaviors or situations, which resemble to games with reference to the principle of rationality,¹⁰⁴ as we saw in respect to Lewis's folk theory in **section 5**. I have argued that the Lewisian conventionalist interpretation of the notion of possibilities in game theory, by introducing the notion of possible worlds and its related notion of 'possibles', advances Lewis's

¹⁰² We take here the notion of rationality as a behavioural scientific principle applied in game theoretical analysis, which provides specific contexts, where the different meanings of the notion of rationality are definable (Rapoport, 1970: 284-285). Lewis formulates an example of conventionalist behavior as a pattern of agents' expected mutual expectations of expectations of desire of acting on the necessary following two conditions: a proposition *s* holds and an act *σ* is performed; which results in at least one agent performing the act *σ*. See figure 45 (Lewis, 1969: 140; 156).

¹⁰³ The aforementioned Lewisian examples of social behavior are based on concepts of two-person games (Rapoport, 1970: 158), which are interpreted as either cooperative or competitive: Lewis (1969: 157) offers such an example from Searle's account of speech acts. As Rapoport (1970: 166) mentions, based on the *n*-person game model, two-person games may have different solutions depending on the followed procedure.

¹⁰⁴ Nonetheless, Rapoport (1970: 285) argues that explanations in sociological terms require the study of departures, such as "systematic pattern-forming departures", from the game theoretical basis and rationality as defined therein.

philosophical study of the systematization of folk common sense.¹⁰⁵ Nevertheless, Lewis's treatment of conventionalism on the basis of the Lewisian interpretation of the principle of 'rationality' does not offer a moral philosophical account. This Lewisian position is evident namely in the example of the problem of deterrence, where Lewis disputes conflating the notions of rationality and morality, or taking them as causally connected.

6. We saw in **subsection 7.3** that Lewis articulates his critique of **utilitarianism**, which is a strand of **consequentialism**, in his treatment of problems of normative ethics and social and political philosophy, such as the tacit treaty of toleration and the practice of deterrence in the penal system. Notably, as we saw in **subsection 7.3**, Lewis does not directly attack the principle itself of utilitarianism, as a normative ethical principle, but problematizes utilitarianist approaches to hypothetical cases in his counterarguments and counterexamples. As mentioned in **section 1**, it can be argued that in his social and political philosophy essays, Lewis follows a method of applied ethics, starting with a case or problem, which has already been situated in a specific discourse or foundational theory, such as utilitarianism and consequentialism, to counter argue principles of the theory itself, by applying philosophico-logical methods, thus offering alternative to the canonical accounts.

Taking the problem of toleration, Lewis (2000 [1989b]: 185-186) examines Mill's defence of the tacit, or explicit and formal, treaty of toleration, counter arguing its Millian basis on the rule of 'neutralism'. Lewis (2000 [1989b]: 178-179) shows that the case of the treaty of conditional toleration, where both sides practice toleration to achieve equilibrium, is utilitarian, yet in fact 'unMillian' (Lewis, 2000 [1989b]: 161). Lewis (2000 [1989b]: 161) explains that, in Millian terms, the defender of toleration does not take sides in the dispute, although he or she may be arguing from factual premises, but not from those, which are part of the dispute itself between suppressors and the suppressed. Lewis (2000 [1989b]: 162) further

¹⁰⁵ One can argue that Lewis (1969, 2000 [1975a]) provides a cognitive model, by appeal to the communicative aspects of social situations bearing resemblances to games, as well as to the possible worlds notion for analyzing possible languages in populations conforming to their own linguistic conventions, which can also be interpreted in game theoretical semantics (Hintikka, 1986). I make this secondary suggestion on the basis of van Benthem's account of three major models (1996: 287-288), which have predominated recent cognitive dynamics in semantics: the model of proofs, which is concerned with variations of deductive proof systems, or formats, determining the truth value of sentences; the model of programs, which is concerned with the computational aspects of natural language and processes of translation; and the model of games, which is concerned with the comparative and evaluative processes involving statements and model structures.

emphasizes that, to argue for toleration from a Millian position, "not only must the factual premises be common ground between the two sides; also a uniform and non-disjunctive argument must be addressed to both". For Lewis (2000 [1989b]: 162-163), a Millian defence of toleration ultimately aims at the persuasion of an audience with diverse opinions, without placing a constraint on the utilitarian principle itself.

In support of his counterargument against Mill's neutralist approach to toleration, Lewis (2000 [1989b]: 163) brings forth the implicit paradox in Mill's account, concerning: i) the protection of "self-regarding conduct" as derived from an ultimate appeal to utility, which is not sufficient to justify the practice of toleration if the utilitarian goal fails; ii) the protection of "other-regarding conduct" as reasoned from the Millian argument of its inseparability from private thought. Besides the evident contradiction in i) between self-regarding conduct by appeal to maximizing utility and the principle of utilitarianism itself, in ii) we encounter a conflict in following classical liberal values, which Mill's account aims to defend, including the public support of our opinions by appeal to reason (Lewis, 2000 [1989b]: 171).¹⁰⁶ Secondly, Lewis (2000 [1989b]: 171) offers additional comments on Marcuse's view of toleration (1970) that, although it may be best that dangerous or false opinions are suppressed, this is not achievable by allowing freedom of expression up to a convenient limit, beyond which suppression is desirable (Lewis, 2000 [1989b]: 172-173).

Taking the problem of the practice of deterrence in the penal system, Lewis's solution of perfect reenactments of criminal cases (2000 [1987]: 238) for deciding on the justness of our penal practices, which is close to a "simulation" (Nolan, 2005:

¹⁰⁶ Gaus (2003: 1-3) identifies the liberal political tradition, such as Mill's account of liberty, primarily with the protection of individual liberties, including the freedoms of speech, thought and conscience, deriving from the devotion to human reason. On the other hand, Machan (1998: 18) argues that there is an inherent to such liberal accounts contradiction between taking individual responsibility for our lives and not acting or formulating plans by our own judgment. In this view, coercive conduct, which has "rights-violating force", is identifiable from the normative framework of freely chosen human action subject to moral evaluation (Machan, 1998: 157-158).

Furthermore, specifically on utilitarianism in view of the liberal tradition, Scalnon (2003: 28), cited in Lewis's essay, comments that traditional utilitarianism has been considered as extreme on the account of assigning rights and liberties, because, on the one hand, "[...] it is maximally specific in the requirements it imposes on an agent, and, since there are no limits to what it may require to be done, it provides a minimum of reliable protection from interference by others". Thus, objections to utilitarianism have been raised, characterizing it as "demanding and intrusive", in favour of other theories of rights, granting "greater discretion" and "better protection" (Scalnon, 2003: 28). However, as Scalnon (2003: 28) points out: "When one individual is given a claim-right or liberty with respect to a certain option, the control that others are able to exercise over their own options is to some degree diminished"; in this way stressing the deficiencies inherent in utilitarian principles.

194-195), offers an obvious methodological appeal to Lewis's modal realism. As Nolan (2005: 194-195) points out, this Lewisian solution proceeds from informal experiments designed on mathematical or other scientific premises, rather than from fictional accounts, or representations, of criminal attempts. I argue that the Lewisian response to the given moral problem is consistent with Lewis's view of the moral subject, as a case of adjustment to any system of belief-desire – say, for instance, utilitarianism – as I discussed in **subsection 7.3**. However, as also seen in **7.3**, Lewis has argued that such a case does not make the system itself correct.

Quinn (2004: 720) interprets Lewis's broader treatment of the problem of penal substitution in two of Lewis's essays (1987; 1997) by virtue of the Christian problem of Atonement. According to Lewis (2000 [1997]: 131), Christians themselves are "double-minded" towards this problem, if they are not ready to argue that "just as Christ paid the debt of punishment owed by all the sinners, so otherwise other innocent volunteers can pay the lesser debts of punishment owed by burglars and murderers". Hence, Lewis dismisses a solution based on the Christian theological tradition: for example, we do not believe that the offender's friend must serve his prison or death sentence, but we accept that he or she can pay the offender's fine,¹⁰⁷ although this logically consists as much in a case of penal substitution as the others (Lewis, 2000 [1997]: 133). Despite leaving the puzzle unsolved, Lewis (2000 [1997]: 135) accepts that in the latter case, there is an indication that:

"[...] both sides agree that penal substitution sometimes makes sense after all, even if none can say how it makes sense. And if both sides agree to that, that is some evidence that somehow they might both be right."

There is some obvious lack of clarity as to whether Lewis's views are broadly consequentialist; for instance, in Lewis's deontic scheme of sphere semantics and its application on problems of metaethics, such as the problem about permission in the example of the hypothetical master and slave game across spheres, which I analyzed in **subsection 7.2**. As I have also shown in **subsection 7.3**, in the moral problem of penal punishment and in the example of act-utilitarians, Lewis further raises a question concerning modal thinking in act-utilitarian approaches, which do not achieve the maximization of expected utility, rather than

¹⁰⁷ In the essay on penal substitution itself, Lewis (1997: 128) argues that common sense morality does not generally abide with the theological and moral notion of penal substitution, but only in certain cases of light punishment, such as paying someone's fine (Lewis, 1997: 134). Lewis (1997: 129-130) also examines the introduction of the compensatory function in the penal system, about which he argues that it is not served by the punishment of criminals, let alone of their innocent substitutes.

a question about the moral principle itself of utilitarianism as a consequentialist moral philosophical theory. Nonetheless, in order to evaluate Lewis's contribution to canonical accounts, his ethics and social philosophy must be considered, in Lewisian modal realist terms, as non-consequentialist. For this purpose, I argue that, although ethics in Lewis's modal metaphysics is broadly construed in deontic logic of obligation and permission, in the modal logic of necessity and possibility, Lewis adopts a critical stance towards consequentialism, which implies a connection between moral values and obligation. As Nolan (2005: 185) mentions:

"[...] while Lewis tells us a story about moral value, it is not explicitly a story about what we ought to do. A natural story is that what we ought to do is maximize moral value, or maximize some balance of moral values. This is even what is suggested by the role of moral value structures in Lewis's semantics for deontic logics (the logics of obligation), but other things Lewis said suggest that he might not endorse this consequentialist thesis about the connection between moral values and obligation."

In consequentialist terms, adding to the encountered paradox of the coincidence of rational and moral concerns, which Lewis wants to examine separately, as we saw earlier in **subsection 7.3**, I argue that Lewis's pluralism, along with his possibilist interpretation of moral considerations and outcomes in problems of ethics, do not comply with a consequentialist interpretation. Evidently, in Lewis's example of the problem of charity, but also in most of his moral problems, apart from the extreme case of nuclear deterrence (Lewis, 2000 [1984b]: 214), Lewis poses problems of normative ethics precisely for delivering a sustained critique of utilitarianism, as a consequentialist moral philosophy supporting the maximization of expected utility. Furthermore, by virtue of his analysis of problems of normative ethics, Lewis brings forth metaethical questions of modality, as we saw in **subsections 7.2 and 7.3**; when, for instance, Lewis considers the application of the principle of toleration as a neutral rule-utilitarian principle, but also when he examines problems of coordination for act-utilitarian agents. The fact that Lewis does not directly question, or criticize utilitarianist principles themselves, besides say 'radical utilitarianism', in canonical ethical terms, does not discredit his critical stance towards utilitarianism and, more generally, towards consequentialism.

9

Conclusion

Following the research objective of analyzing and evaluating the contributions of Lewis's work in the area of ethics and social philosophy, I have shown that Lewis positions his work more or less systematically, and certainly consistently, in respect to several canonical accounts of metaethics, ethics, and social and political philosophy. Moreover, I have shown that Lewis debates these accounts by counterarguing select principles of canonical accounts and by offering counterexamples to normative ethical problems. Alongside with positioning Lewis's ethics and social philosophy in respect to canonical accounts, I have argued that Lewisian modal realism is indispensable to the above. For this reason, I have placed emphasis on Lewis's modal realist interpretations of philosophical problems, in response to canonical views, but also taking into consideration several modal realist counterarguments.

In view of the above, I now summarize several substantial points concerning my evaluative conclusions presented in **sections 7 and 8** as follows:

- Lewis does not counteract the nomological deductive interpretation of problems of determinism and non-determinism. Yet he departs from this basic interpretation in differentiating between causally dependent relations and nomological causality and in accounting for so-called nomic 'divergences' of events from the basic model. Additionally, Lewis disputes *sui generis* definitions of truths, as pertaining, for instance, to natural laws.
- Lewis introduces the problem of free will with his formulation of the 'compatibility premise' and his interpretation of compatibilism and incompatibilism to be determined by one's character. These innovations allow Lewis to restate problems of determinism and non-determinism by bypassing causal chains, in this way also resisting consequentialism.

- Lewis's dispositional value theory on the basis of 'the desire to desire' offers an answer to questions of metaethics, such as what is moral value. Lewis claims to offer an 'irrealist' account of value, arguing that values loosely exist and in relation to belief-desire, which is defined by one's own place at a world, taking value to be contingent with respect to this or that world. However, Lewis's lack of regular and explicit argumentation for realism against anti-realism or idealism does not simply imply that his account of moral value does not presume the realism of value.
- Lewis makes propositions for virtue ethics, which do not contradict Lewis's position on the non-essentialism of properties – aside from say the property of being human – which Lewis supports in his value pluralism in modal metaphysics. Lewis argues that moral properties can be ascribed or self-ascribed to world-bound individuals by reference and according to a world; yet for Lewis, there are no necessary values *per se*.
- Lewis proposes schemes for deontic ethics, as evaluative value structures and as formal propositions of deontic logic. Still he stresses that such schemes can be construed on any given moral philosophical principle, thus they do not necessarily motivate or result in good actions, which Lewis differentiates from right actions. Additionally, Lewis's application of deontic conditionals on his system of sphere semantics shows that Lewis's metaphysical pluralism manifests pluralistic principles, such as the conditionality and relativity of value, which Lewis endorses in his value theory.
- Lewis also offers a conventionalist account in his treatment of conventionalism on the basis of his interpretation of the principle of 'rationality'. Lewis's interpretation of convention stresses the role of salience in selecting conventions in populations, thus providing explanations for folk moral psychological and other social behavior. However, it does not offer a systematic moral philosophical account.
- Lewis offers a sustained critique of the moral philosophical account of utilitarianism and its associated strand of consequentialism. Lewis treats problems of normative ethics, such as the tacit treaty of toleration and the penal practice of deterrence, in response to a range of utilitarianist accounts. However, Lewis's critique brings forth mainly questions of modality on how to maximize value, rather than disputing utilitarian principles themselves.

In pursuing my research objective and reaching the aforementioned conclusions, I have also discussed Lewis's defense of the Humean versus the anti-Humean thesis of belief as desire, where Lewis's value theory rests, bearing also an influence upon Lewis's accounts of folk 'moral science' and of value pluralism. The Lewisian value theory schematic, but nonetheless systematic, approach supports Lewis's advancement of the Humean thesis, which underpins Lewis's treatment of problems of ethics and social philosophy. In my view, this defense is also indicative of Lewis's philosophical efforts to systematize folk moral views with his dispositional value theory, but also his conventionalist and modal metaphysical interpretations.

Clearly and aside other concerns, Lewis's modal realist appeal, also in his work on ethics and social philosophy, is methodological. Taken as such, it makes sense that, as Hall (2010) suggests, although Lewis's philosophical method is not "bottom up", it is best if Lewis's philosophical inquiry and relevant results are presented "bottom up". Hence, my aim to offer an interpretation of the canonization of Lewis's ethics and social philosophy justifies my analysis of Lewis's folk ethics, in its conventionalist and modal metaphysical approaches, and in view of Lewis's pluralism. Applying possible worlds interpretations on Lewis's folk ethics, besides offering an additional level of analysis, has opened up new modes of systematic interpretations of folk psychology, which Lewis himself adopted alongside his dispositional value theory.

Finally, following Nolan's claim (2005: 228) that Lewis's philosophical system is not "a monolith that must be taken or left as a whole", I would like to explain my take on the philosopher's overall project by referring to Lewis's own suggestion that:

"[...] the skeptical reader will consider breaking up the package and taking the parts that suit him, but I have not done all I could have done to make his task easy. For, after all, my principal interest has not been to proselytize but to figure out what I should believe."

With this in mind, I have aimed at bringing to light Lewis's ethics and social philosophy as a part of "the package", in my belief that it will be influential on philosophy, but also beyond the academy.

Bibliography

- Allen, Colin and Bekoff, Marc, *Species of Mind: The Philosophy and Biology of Cognitive Ethology*, Cambridge, MA: The MIT Press, 1997.
- Allen, Colin and Bekoff, Marc, "Animal Minds, Cognitive Ethology and Ethics". In *The Journal of Ethics*, September 2007, 11, pp. 299-317.
- Anscombe, G.E.M., *The Collected Philosophical Papers of G.E.M. Anscombe, Vol 2: Metaphysics and the Philosophy of Mind*, Oxford: Blackwell, 1981.
- Bennett, Jonathan, *Rationality: An Essay Towards An Analysis*, London: Routledge, 1964.
- Berker, Selim, "The Unity of Grounding", draft paper, Harvard University, 4 August 2015.
Available at: <https://scholar.harvard.edu/sberker/publications/unity-grounding>
(last accessed on 20/9/2019)
- Berto, Francesco and Plebani, Matteo, *Ontology and Metaontology: A Contemporary Guide*, London and New York: Bloomsbury, 2015.
- Blackburn, Simon, *Spreading the Word: Groundings in the Philosophy of Language*, Oxford: Clarendon Press, 1984.
- Blackburn, Simon, "Just Causes". In *Philosophical Studies*, 61, 1991, pp. 3-17.
- Blackburn, Simon, "Quasi-Realism no Fictionalism". In M. E. Kalderon (ed.), *Fictionalism in Metaphysics*, Oxford: Clarendon Press, 2005, pp. 320-338.
- Blackburn, Simon, "The Semantics of Non-Factualism, Non-Cognitivism and Quasi-Realism". In *The Blackwell Guide to the Philosophy of Language*, M. Devitt and R. Hanley (eds.), Malden, MA: Blackwell, 2006, pp. 244-251.
- Brandom, Robert B., *From Empiricism to Expressivism: Brandom Reads Sellars*, Cambridge, MA: Harvard University Press, 2014.
- Buehler R.J., "Counterfactuals: Notes Based on Counterfactuals by David Lewis", Berkeley: Berkeley University, 2014.
Available at: <https://math.berkeley.edu/~buehler/Counterfactuals Notes.pdf>
(last accessed on 10/6/2018)
- Burge, Tyler, "Philosophy of Language and Mind: 1950-1990". In *The Philosophical Review*, Vol. 101, No 1, Jan. 1992, pp. 3-51.
- Carnap, Rudolf, "Value Concepts". In *Synthese* 194, 2017 (1958), pp. 185-194.

- Chihara, Charles, *The Worlds of Possibility: Modal Realism and the Semantics of Modal Logic*, Oxford: Clarendon Press, 1998.
- Chisholm, Roderick, "Identity through possible worlds: some questions". In *Noûs*, Vol. 1, No 1, Mar. 1967, pp. 1-8.
- Cocchiarella, Nino B. and Freund, Max, *Modal Logic: An Introduction to Its Syntax and Semantics*, Oxford: Oxford University Press, 2008.
- Copp, David, "Introduction: Metaethics and Normative Ethics". In D. Copp (ed.) *Oxford Handbook of Ethical Theory*, Oxford: Oxford University Press, 2006, pp. 3-35.
- Davidson, Donald, "Radical Interpretation". In *Dialectica*, Vol. 27, 1973, pp. 313-328.
- Davidson, Donald, *Inquiries into Truth and Interpretation*, Oxford: Clarendon Press, 1984.
- Divers, John, *Possible Worlds*, London and New York: Routledge, 2002.
- Divers, John, "The Modal Metaphysics of Alvin Plantinga". In D. Baker (ed.), *Alvin Plantinga*, Cambridge: Cambridge University Press, 2007, pp. 71-92.
- Edgington, Dorothy, "On Conditionals". In D. M. Gabbay and F. Guenther, *Handbook of Philosophical Logic*, Vol. 14, Dordrecht: Springer, 2007, pp. 127-221.
- Fitting, Melvin and Mendelsohn, Richard, *First Order Modal Logic*, Synthese Library, Vol. 277, New York: Lehman College and CUNY, 1998.
- Gamut, L.T.F., *Logic, Language, and Meaning: Intentional Logic and Logical Grammar*, Vol. II, Chicago and London: The University of Chicago Press, 1991.
- Garson, James W., *Modal Logic for Philosophers*, Cambridge: Cambridge University Press, 2006.
- Gaus, Gerald F., *Contemporary Theories of Liberalism: Public Reason as a Post-Enlightenment Project*, London: Sage, 2003.
- Gauthier, David, "Deterrence, Maximization and Rationality". In D. MacLean (ed.), *The Security Gamble: Deterrence in the Nuclear Age*, Totowa, New Jersey: Rowman and Allanheld, 1984, pp. 474-495.
- Hacking, Ian, *The Emergence of Probability: A Philosophical Study of Early Ideas About Probability, Induction and Statistical Inference*, Cambridge: Cambridge University Press, 2006.
- Hawthorne, John, "A Note on Languages and Language". In *Australasian Journal of Philosophy* 68, 1990, pp. 116-118.

- Hintikka, Jaakko, "Game-Theoretical Semantics as a Synthesis of Verificationist and Truth-Conditional Meaning Theories". In *New Directions in Semantics*, E. LePore (ed.), London: Academic Press Ltd., 1986, pp. 235-258.
- Hume, David, *An Inquiry Concerning Human Understanding*, P. Millican (ed.), Oxford: Oxford University Press, 2007.
- Hume, David, *Essays Moral, Political and Literary*, Vol. I and II, 1742-1752.
Available at: www.hetwebsite.net/het/texts/hume/humeessays.htm
(last accessed on: 1.5.2019)
- Jamieson, Dale, "David Lewis on Convention". In *Canadian Journal of Philosophy*, Vol 5, No 1, September 1975, pp. 73-81.
- Jolley, Nicholas, *Leibniz*, London and New York: Routledge, 2005.
- Kekes, John, *The Morality of Pluralism*, Princeton, NJ: Princeton University Press, 1993.
- Knowles, Dudley, *Political Philosophy*, London: Routledge, 2001.
- Kripke, Saul, *Naming and Necessity*, Oxford: Blackwell, 1980.
- Lewis, David K., "Counterpart Theory and Quantified Modal Logic". In *The Journal of Philosophy*, Vol. 65, no. 5, Mar. 7, 1968, pp. 113-126.
- Lewis, David K., *Convention: A Philosophical Study*, Harvard University Press: Cambridge, MA, 1969.
- Lewis, David K., "How to Define Theoretical Terms". In *The Journal of Philosophy*, Vol. LXVII, No. 13, (Jul. 9, 1970), pp. 427-446.
- Lewis, David K., "Causation". In *The Journal of Philosophy*, Vol. 70, No. 17, Oct. 11 1973a, pp. 556-567.
- Lewis, David K., *Counterfactuals*, Oxford: Basil Blackwell, 1973b.
- Lewis, David K., "Radical Interpretation". In *Synthese*, Vol. 23, 1974a, pp. 331-344.
- Lewis, David K., "Attitudes *De Dicto* and *De Se*". In *The Philosophical Review*, Vol. 88, No 4 (Oct. 1979a), pp. 513-543.
- Lewis, David K., "Counterfactual Dependence and Time's Arrow". In *Noûs*, Vol. 13, No. 4, Special Issue on Counterfactuals and Laws, (Nov. 1979b), pp. 455-476.
- Lewis, David K., "Individuation by Acquaintance and by Stipulation". In *The Philosophical Review*, XCII, No. 1 (Jan. 1983a), pp. 3-32.
- Lewis, David K., "New Work for a Theory of Universals". In *Australasian Journal of Philosophy*, Vol. 61, No 4, 1983b, pp. 343-377.

- Lewis, David K., "Introduction". In *Philosophical Papers Vol. I*, New York, Oxford: Oxford University Press, 1983c, pp. ix-xii.
- Lewis, David K., "General Semantics". In *Philosophical Papers Vol. I*, New York, Oxford: Oxford University Press, 1983d (1972), pp. 189-232.
- Lewis, David K., "Languages and Language". In D. K. Lewis, *Philosophical Papers Vol. I*, New York, Oxford: Oxford University Press, 1983e (1975a), pp. 163-188.
- Lewis, David K., "Truth in Fiction". In D. K. Lewis, *Philosophical Papers Vol. I*, New York, Oxford: Oxford University Press, 1983f (1978), pp. 261-275.
- Lewis, David K., "Putnam's Paradox". In *Australasian Journal of Philosophy*, Vol. 62, No. 3, Sept. 1984a, pp. 221-236.
- Lewis, David K., *On The Plurality of Worlds*, New York: Blackwell, 1986a.
- Lewis, David K., "Counterfactuals and Comparative Possibility". In D. K. Lewis, *Philosophical Papers Vol. II*, New York, Oxford: Oxford University Press, 1986b (1973), pp. 3-31.
- Lewis, David K., "Utilitarianism and Truthfulness". In D. K. Lewis, *Philosophical Papers Vol. II*, New York, Oxford: Oxford University Press, 1986c (1972), pp. 340-342.
- Lewis, David K., "Probabilities of Conditionals and Conditional Probabilities". In *Philosophical Papers Vol. II*, Oxford: Oxford University Press, 1986d (1976), pp. 133-158.
- Lewis, David K., "Are We Free to Break the Laws?" In *Philosophical Papers Vol. II*, New York, Oxford: Oxford University Press, 1986e (1981a), pp. 291-298.
- Lewis, David K., "Noneism or Allism?" In *Mind*, New Series, Vol. 99, No. 393, (Jan. 1990), pp. 23-31.
- Lewis, David K., "Elusive Knowledge". In *Australasian Journal of Philosophy*, Vol. 74, No. 4, (Dec. 1996a), pp. 549-567.
- Lewis, David K., "Scriven on Human Unpredictability". In D. K. Lewis, *Papers in Ethics and Social Philosophy*, Cambridge: Cambridge University Press, 2000 (1966), pp. 244-250.
- Lewis, David K., "Semantic Analyses for Dyadic Deontic Logic". In D. K. Lewis, *Papers in Ethics and Social Philosophy*, Cambridge: Cambridge University Press, 2000 (1974b), pp. 5-19.

- Lewis, David K., "A Problem About Permission". In D. K. Lewis, *Papers in Ethics and Social Philosophy*, Cambridge: Cambridge University Press, 2000 (1975b), pp. 20-33.
- Lewis, David K., "Convention: Reply to Jamieson". In D. K. Lewis, *Papers in Ethics and Social Philosophy*, Cambridge: Cambridge University Press, 2000 (1976b), pp. 136-144.
- Lewis, David K., "Reply to McMichael". In D. K. Lewis, *Papers in Ethics and Social Philosophy*, Cambridge: Cambridge University Press, 2000 (1978), pp. 34-36.
- Lewis, David K., "Why Ain'cha Rich?". In D. K. Lewis, *Papers in Ethics and Social Philosophy*, Cambridge: Cambridge University Press, 2000 (1981b), pp. 37-41.
- Lewis, David K., "Devil's Bargains and the Real World". In D. K. Lewis, *Papers in Ethics and Social Philosophy*, Cambridge: Cambridge University Press, 2000 (1984b), pp. 201-218.
- Lewis, David K., "The Punishment that Leaves Something to Chance". In D. K. Lewis, *Papers in Ethics and Social Philosophy*, Cambridge: Cambridge University Press, 2000 (1987), pp. 227-243.
- Lewis, David K., "Desire as Belief". In D. K. Lewis, *Papers in Ethics and Social Philosophy*, Cambridge: Cambridge University Press, 2000 (1988), pp. 42-54.
- Lewis, David K., Johnston, Mark, and Smith, Michael, "Dispositional Theories of Value". In D. K. Lewis, *Papers in Ethics and Social Philosophy*, Cambridge: Cambridge University Press, 2000 (1989a), pp. 68-94.
- Lewis, David K., "Mill and Milquetoast". In D. K. Lewis, *Papers in Ethics and Social Philosophy*, Cambridge: Cambridge University Press, 2000 (1989b), pp. 159-186.
- Lewis, David K., "Meaning Without Use: Reply to Hawthorne". In D. K. Lewis, *Papers in Ethics and Social Philosophy*, Cambridge: Cambridge University Press, 2000 (1992), pp. 145-151.
- Lewis, David K., "Evil for Freedom's Sake?". In D. K. Lewis, *Papers in Ethics and Social Philosophy*, Cambridge: Cambridge University Press, 2000 (1993), pp. 101-127.
- Lewis, David K., "Desire as Belief II". In D. K. Lewis, *Papers in Ethics and Social Philosophy*, Cambridge: Cambridge University Press, 2000 (1996b), pp. 55-67.
- Lewis, David K., "Illusory Innocence?". In D. K. Lewis, *Papers in Ethics and Social Philosophy*, Cambridge: Cambridge University Press, 2000 (1996c), pp. 152-158.

- Lewis, David K., "Do We Believe in Penal Substitution?". In D. K. Lewis, *Papers in Ethics and Social Philosophy*, Cambridge: Cambridge University Press, 2000 (1997), pp. 128-135.
- Lewis, David K., "Tensed Quantifiers". In D. W. Zimmerman (ed.), *Oxford Studies in Metaphysics*, Vol. 1, Oxford: Oxford University Press, 2002 (2004), pp. 3-14.
- Lewis, David K., "Quasi-Realism is Fictionalism". In M. E. Kalderon (ed.), *Fictionalism in Metaphysics*, Oxford: Clarendon Press, 2005, pp. 314-321.
- Lycan, William G., "Pot Bites Kettle: A Reply to Miller". In *Australasian Journal of Philosophy*, Vol. 69, No. 2, (June 1991), pp. 212-213.
- Machan, Tibor, *Classical Individualism: The Supreme Importance of Each Human Being*, London and New York: Routledge, 1998.
- Marcuse, Herbert, "Repressive Tolerance" with postscript. In R. P. Wolff, B. Moore, H. Marcuse, *A Critique of Pure Tolerance*, Boston: Beacon Press, (1965, 1969) 1970, pp. 81-123.
- McMichael, Alan, "Too Much of A Good Thing: A Problem in Deontic Logic". In *Analysis*, Vol. 38, No 2, (Mar. 1978), pp. 83-84.
- Mill, John S., *On Liberty*, London: Watts and Co, 1929.
- Miller, Alexander, *An Introduction to Contemporary Metaethics*, Cambridge: Polity Press, 2003.
- Miller, Richard B., "Genuine Modal Realism: Still the Only Non-Circular Game in Town". In *Australasian Journal of Philosophy*, Vol. 71, No. 2, June 1993, pp. 159-160.
- Moses, Yoram, "Reasoning About Knowledge and Belief". In F. van Harmelen, V. Lifschitz and B. Porter (eds.), *Handbook of Knowledge Representation*, Amsterdam: The Netherlands: Elsevier, 2008.
- Nolan, Daniel, "Three Problems for 'Strong' Modal Fictionalism". In *Philosophical Studies* 87, 1997, pp. 259-275.
- Nolan, Daniel, *David Lewis*, Chesam: Acumen Publishing, 2005.
- Norris, Christopher, "Will the real Saul Kripke please stand up? Fiction, philosophy and possible worlds". In *Textual Practice* 17(2), 2003, pp. 225-251.
- Nozick, Robert, "Newcomb's Problem and Two Principles of Choice". In N. Rescher et al. (eds.), *Essays in Honor of Carl G. Hempel*, Dordrecht: D. Reidel Publishing Co and Springer, 1969, pp. 114-146.
- Osborne, Martin and Rubinstein, Ariel, *A Course in Game Theory*, Cambridge MA: The MIT Press, 1994.

- Plantinga, Alvin, *God, Freedom and Evil*, New York: Harper and Row, 1974.
- Plantinga, Alvin, *Essays in the Metaphysics of Modality*, M. Davidson (ed.), Oxford: Oxford University Press, 2003.
- Prior, Arthur N. and Fine, Kit, *Worlds, Times and Selves*, London: The Trinity Press, 1977.
- Psillos, Stathis, "Past and Contemporary Perspectives on Explanation". In Theo Kuipers (ed.), *Handbook of the Philosophy of Science: Focal Issues*, Amsterdam: Elsevier BV, 2007, pp. 121-197.
- Putnam, Hilary, *Realism and Reason: Philosophical Papers, Vol. 3*, Cambridge: Cambridge University Press, 1983.
- Putnam, Hilary, *Realism with a Human Face*, James Conant (ed.), Cambridge, MA: Harvard University Press, 1990.
- Quine, Willard van Orman, "Existence and Quantification". In W.V.O. Quine, *Ontological Relativity and Other Essays*, New York: Columbia University Press, 1969, pp. 91-113.
- Quine, Willard van Orman, "Worlds Away". In *The Journal of Philosophy: Volume LXXIII*, No. 22, December 16, 1976, pp. 859-863.
- Quinn, Philip L., "Papers in Ethics and Social Philosophy by David Lewis". In *Noûs*, Vol. 38, No. 4, (Dec. 2004), 2004, pp. 711-730.
- Rapoport, Anatol, *N-person Game Theory: Concepts and Applications*, Ann Arbor: The University of Michigan Press, 1970.
- Rawls, John, *The Law of Peoples*, Cambridge, MA: Harvard University Press, 1999.
- Rousseau, Jean Jacques, *A Discourse on Inequality*, London: Penguin Classics, 1984.
- Scalnon, T. M., *The Difficulty of Tolerance: Essays in Political Philosophy*, Cambridge: Cambridge University Press, 2003.
- Schiffer, Stephen, "Paradox and the A Priori". In T. S. Gendler and J. Hawthorne (eds.), *Oxford Studies in Epistemology*, Vol. I, Oxford: Clarendon Press, 2005, pp. 273-310.
- Scriven, Michael, "An Essential Unpredictability in Human Behavior". In B. Wolman and E. Nagel (eds.), *Scientific Psychology: Principles and Approaches*, New York: Basic Books, 1965, pp. 411-425.
- Searle, John, "What is Language: Some Preliminary Remarks". In *Ethics & Politics*, XI, 2009, 1, pp. 173-202.

- Sellars, Wilfrid, *Pure Pragmatics and Possible Worlds: The Early Essays of Wilfrid Sellars*, J. Sicha (ed.), Atascadero: Ridgeview Publishing Company, 1980.
- Sider, Theodore, "Ontological Realism". In D. Chalmers, D. Manley, and R. Wasserman (eds.), *Metametaphysics*, Oxford: Oxford University Press, 2009, pp. 384-423.
- Stalnaker, Robert, "Possible Worlds". In *Noûs*, Vol. 10, No 1, 1976, pp. 65-75.
- Stalnaker, Robert, "Notes on Conditional Semantics", Department of Linguistics and Philosophy, Cambridge, MA: The MIT Press, 1990.
- Staudacher, Marc, "Between Conventions and Social Norms", ILLC Dissertation Series, Amsterdam: University of Amsterdam, 2010.
Available at: www.illc.uva.nl/Research/Publications/Dissertations/DS-2010-07.text.pdf (last accessed on: 4/7/ 2018)
- Van Benthem, Johan, *Exploring Logical Dynamics*, Stanford, CA: CSLI Publications, 1996.
- Van Inwagen, Peter and Zimmerman, Dean, "Introduction". In P. van Inwagen and D. Zimmerman (eds.), *Metaphysics: The Big Questions*, Oxford: Blackwell, 2008, pp. 1-13.
- Van Inwagen, Peter, "The Consequence Argument". In P. van Inwagen and D. Zimmerman (eds.), *Metaphysics: The Big Questions*, Oxford: Blackwell, 2008, pp. 450-456.
- Van Rooij, Robert and Schulz, Katrin, "Non-monotonic Reasoning in Interpretation". In J. van Benthem and A. T. Meulen (eds.), *Handbook of Logic and Language*, Amsterdam: Elsevier, 2010, pp. 839-856.
- Unger, Peter, "The Mental Problems of the Many". In D. W. Zimmermann (ed.), *Oxford Studies in Metaphysics*, Vol. 1, Oxford: Oxford University Press, 2004, pp. 195-222.

Lewis's original publications in ethics and social philosophy

- "Scriven on Human Unpredictability: With Jane S. Richardson". In *Philosophical Studies* 17, 1966, pp. 69-74.
- "Utilitarianism and Truthfulness". In *Australasian Journal of Philosophy* 50, 1972, pp. 17-19.

- "Semantic Analyses for Dyadic Deontic Logic". In S. Stenlund (ed.), *Logical Theory and Semantic Analysis: Essays Dedicated to Stig Kanger on his Fiftieth Birthday*, Dordrecht: Reidel, 1974, pp. 1-14.
- "A Problem About Permission". In E. Saarinen et al (eds.), *Essays in Honour of Jaakko Hintikka*, Vol. 124, Dordrecht: Reidel, 1975, pp. 163-176.
- "Convention: Reply to Jamieson". In *The Canadian Journal of Philosophy* 6, 1976, pp. 113-120.
- "Reply to McMichael". In *Analysis* 38, 1978, pp. 85-86.
- "Are We Free to Break the Laws?". In *Theoria* 47, 1981, pp. 113-121.
- "Why Ain'cha Rich?". In *Noûs* 15, 1981, pp. 377-380.
- "Devil's Bargains and the Real World". In D. MacLean (ed.), *The Security Gamble: Deterrence Dilemmas in the Nuclear Age*, Totowa, New Jersey: Rowman and Allanheld, 1984 (pages unknown).
- "Buy Like a MADman, use like a NUT". In *QQ* 6, No. 2, 1986, pp. 5-8.
- "The Punishment that Leaves Something to Chance". In *Proceedings of the Russellian Society* (University of Sydney), 12, 1987, pp. 53-67; and in *Philosophy and Public Affairs* 18, 1989, pp. 53-67.
- "Desire as Belief". In *Mind* 97, 1988, pp. 323-332.
- "The Trap's Dilemma". In *The Australasian Journal of Philosophy* 66, 1988, pp. 220-223.
- "Dispositional Theories of Value". In *The Proceedings of the Aristotelian Society*, Supplementary Volume 63, 1989, pp. 113-137.
- "Mill and Milquetoast". In *The Australasian Journal of Philosophy* 67, 1989, pp. 152-171.
- "Academic Appointments: Why Ignore the Advantage of Being Right?". In *Ormond Papers*, Ormond College, University of Melbourne, 1989 (pages unknown).
- "Meaning Without Use: Reply to Hawthorne". In *The Australasian Journal of Philosophy* 70, 1992, pp. 106-110.
- "Evil for Freedom's Sake?". In *Philosophical Papers* 22, 1993, pp. 149-172.
- "Desire as Belief II". In *Mind* 105, 1996, pp. 303-313.
- "Illusory Innocence?". In *Eureka Street* 5, 1996, pp. 35-36.
- "Do We Believe in Penal Substitution?". In *Philosophical Papers* 26, 1997, pp. 203-209.
- "Quasi-Realism is Fictionalism". In *Fictionalism in Metaphysics*, M. E. Kalderon (ed.), Oxford: Clarendon Press, 2005, pp. 314-321.

Appendixes

Appendix I

Objections and Replies

I. Objection: Lewisian modal realism does not solve the common analytic problems of the Humean supervenience thesis and of individuation.

Reply: Lewis offers solutions on these problems drawing from his folk moral science, which, as we saw, is a combination of folk psychology and folk ethics, by concentrating on comparisons between language users in a population and cross-world comparisons. Characteristically for the modal realist account – and contrary to haecceitist accounts – where there is a distinction of possibilities without qualitative difference of worlds,¹⁰⁸ the modal realist sees different possible individuals at the same world (Lewis, 1986a: 243). We saw that the Lewisian modal realist approaches to possible worlds and individuals are not only useful for the analysis of thought and language, but also for the analysis of social practices of language usage; as Lewis (1986a: 40) explains:

“Suppose we want a systematic grammar, covering not only syntax but semantics, for a natural language or some reasonable imitation or fragment thereof. Such a grammar is meant to plug into an account of the social practice of using language. It encapsulates the part of the account that is different for different linguistic communities who are party to different conventions of language.” Specifically on the given problems of:

i) the Humean supervenience thesis, Lewis offers the above explanation for semantics in a population by introducing the question of the Humean psychophysical supervenience thesis, which supports the view that mental facts are ontologically dependent on physical facts, following from them by necessity (Kripke, 1980: 155). Lewis (1986a: 14-17) shows that the substitution of the natural language modal ‘*could*’ for the diamond, in modal realist terms, creates distortion with regard to the possible worlds analysis of this thesis, which supports that there *could* be no difference of the mental and physical sorts. Lewis (1986a: 15) explains that when the diamond, the modal operator ‘*possibly*’, is used in place of the modal ‘*could*’, the narrow supervenience thesis is translated as:

¹⁰⁸ It can be argued here that Lewis refers to cases of isomorphism, which characterizes the strand of modal realism abiding to the view that pictorial descriptions stand for possible worlds.

(1) "It is not the case that, possibly, there are two things which have a difference of the one sort without any difference of the other sort."

Or, in possible worlds terms:

(2) "It is not the case that there is some world *W* such that, at *W*, two things have a difference of the one sort but not the other."

For Lewis (1986a: 16), statement (2) is only problematic when the modal operator '*possibly*' is applied in the broad supervenience thesis translated as:

(3) "There could be no mental difference between two people without there being some physical difference, whether intrinsic or extrinsic between them."

Or, in possible worlds terms:

(4) "There is no world wherein two people differ mentally without there being some physical difference, intrinsic or extrinsic, between them."

Statement (4) means that:

(5) "There could be no mental differences without some physical difference of the sort that could arise between people in the same world."

Statement (5) is not what the broad supervenience thesis states and disqualifies us from making cross-world comparisons, because we are ignoring the extrinsic physical differences arising between people in different worlds.¹⁰⁹

We saw that in the social realm of folk 'moral science', Lewis (1969: 195-202) shows his interest in the use of language as semantics in a population also in *de re* modality as belief-desire, which allows him to continue to critically examine the psychophysical supervenience thesis, without dismissing the materialist position, or his subjective and naturalistically metaphysical dispositional value theory.¹¹⁰ Lewis (1986a: 223) rather supports the view of the subject's adjustment to systems of beliefs and desires, for which he offers another example from folk psychology. For example, an individual, named Fred, has conflicting feelings of love and hate towards his father. In this case, there is no right and wrong answer as to whether the physical facts determine the mental facts of Fred's dual system of belief-desire; there is no "one or other eligible assignment of content" that best fits the physical facts, rather two systems fit equally well and according to one system Fred loves and according to the other he hates. For Lewis (1986a: 223), what the physical facts

¹⁰⁹ Nolan (2005: 29) interprets Lewis's view as a partial defence of the Humean supervenience thesis, but not as an endorsement of the supervenience thesis as the truth.

¹¹⁰ For instance, Lewis discusses this thesis through the Moorean argument of intrinsic value and intrinsic nature; see (Lewis, 1986a: 195; 2000b [1989]: 71). For commentary on further references to Moore by Lewis, see also (Nolan, 2005: 208-209).

verify is not Fred's feelings of love and hate, but his state of indeterminacy.

ii) Lewis uses the term individuation in his dispositional theory of value, by applying *de re*, *de dicto*, and *de se* analysis to analyze the dispositions to value of individuals. We saw that Lewis supports a moderate realist, or as he calls it 'irrealist', view regarding the ontological status of moral values. When located in folk theory, Lewis's (2005: 317) view of moral dispositions is rather collective and world specific, than traditionally naturalistic. Therefore, Lewis (2005: 316) recognizes that "the diversity we observe might be due, not to differences in our underlying dispositions, but rather to the fact that different ones of us have actualized different ones of our shared dispositions". Since valuing is desiring to desire, one can desire *de dicto* or *de se*, depending on what they are disposed to value, but they must distinguish between values *de dicto* and *de se* (Lewis, 2000 [1989a]: 75).

Returning to Fred's example, apart from his contradictory feelings towards his father, Fred has other desires, or *attitudes* (Lewis, 1986a: 125), which can be characterised as *de se*, or egocentric, or "essentially indexical": for instance, to be virtuous, or magnanimous, or honourable. What this means, for Lewis (2000 [1989a]: 73), is that Fred wants to realise for himself one of a certain 'property', or class of possibilities for an individual; this class, which is also an 'egocentric proposition', offers the content of Fred's *de se* desire. Fred may also have desires that the world realizes one of a certain class of possibilities for this world, which is a *de dicto* proposition and gives the content of his desire, as well as his desire for the proposition to be true (Lewis, 2000 [1989a]: 73).

Another Lewisian folk psychological example about individuation in *de se* modality (1986a: 198) clarifies Lewis's non-traditionally naturalistic view of his dispositional value theory, as a sort of analytic naturalism, while it highlights his criticism of the supervenience thesis. We saw that in Lewisian possible worlds, individuals are taken together with their surroundings. Take Hubert Humphrey, who is a candidate in upcoming elections. In *de re* representation, Humphrey having a property according to a world means that Humphrey himself having that very property is part of that world (Lewis, 1986a: 200). When transworld identity works by *de re* representation, where there is an overlap between worlds, "the common part of two worlds is supposed to have different properties in one world and in the other" (Lewis, 1986a: 199). There is the difficulty of solving the problem of

a contradiction concerning Humphrey's individuation: for instance, Humphrey having nine toes on his feet at one world, where he has a common part, and ten at another world, where his common part also exists; so, how many toes does the individual Humphrey have? Conversely, as Lewis (1986a: 200) explains, in his genuine modal realism, "the whole of Humphrey is part of different worlds, with different properties at different ones",¹¹¹ therefore, there is no logical contradiction concerning Humphrey's properties.

Clearly, for Lewis (1986a: 219-220), transworld individuation encounters a problem with folk psychology: "A modal realism which makes ordinary things to be transworld individuals disagrees gratuitously with common opinion". In modal metaphysical terms, Lewis (1986a: 219) distinguishes between the perdurantist view of a person, perduring through momentary stages, and the transworld individual perduring across worlds,¹¹² in consistence with the Lewisian *de se* modality, which combines *de re* and *de dicto* representation to capture the content of one's belief-desire. Since, in *de se* modality, the belief-desire concerns an individual's own place at a world of which the individual is wholly part (Lewis, 1986a: 211), where one takes the view of an overlap between worlds in transworld individuation, it is not possible to analyze one's belief-desire in terms of *de se* modality. Moreover, for Lewis (1986a: 219), "the supposed transworld person, no matter how well unified by counterpart relations is not the sort of integrated self that is capable of self-interest". Consequently, in Lewisian genuine modal realist terms, there is a world where Humphrey waves and wins the elections and there is another world where he doesn't; yet there is no overlapping world, where Humphrey both wins and loses, while we are not able to capture his belief-desire in *de se* representation.

¹¹¹ Additionally, Lewis (1986a: 202) considers a mereological analysis of transworld identity, in the case of the vast transworld individual Humphrey, who is composed of "distinct parts from different worlds", as well as of transworld identity in linguistic ersatzism, in the case of Humphrey, whose intrinsic properties are presented correctly, or misrepresented, in the press, to conclude that these transworld identifications do not pose a logical contradiction. See also Lewis (1986a: 216), for a description of the Lewisian view of counterpart existence based on the example of Humphrey.

¹¹² In analytic metaphysics, perdurantism is the claim that objects or individuals persist, in terms of their existence through time, by means of different temporal parts or stages. Lewis (1986a: 202) explains that, "something persists iff, somehow or other it exists at various times; this is the neutral word. Something perdures iff, it persists by having different temporal parts, or stages, at different times, though no one part of it is wholly present at more than one time; whereas it endures iff it persists by being wholly present at more than one time". Lewis (1990: 31) does not take problems of existence in modal metaphysics as problems amounting to a substantive thesis on what exists of what there is, but only with regard to the meaning of the word 'exist'.

2. Objection: Lewisian modal realism does not offer an answer to the analytic question of indeterminacy in folk theory.

Reply: Insisting on the use of theoretical terms for his analysis of folk theory, Lewis (2000 [1989a]: 92) recognizes that his account of folk 'moral science', as any science, submits to creating indeterminacy and equivocal analyticity, by introducing names for candidates that do not deserve, or only imperfectly deserve the names, including the term or name 'value'. Resolutions of the indeterminacy, either make the term's indeterminacy analytic, that the term must be frame-independent, or that it must be an equivalence relation; however, not both at once, because this brings an error to the term's theory (Lewis, 2000 [1989a]: 92).

Drawn from Lewis's conventionalist account (2000 [1989a]: 75), another example emphasizes the indeterminacy of folk moral psychological platitudes, such as 'egoism', which is the thesis that one's own happiness is the only value. In *de dicto* representation, egoism says that "for each person Y, the proposition that Y is happy is the only value", which is inconsistent; yet *de se* representation says that "the property of happiness", that is the egocentric proposition that one is happy, "is the only value" (Lewis, 2000 [1989a]: 75), which is not inconsistent. What is alleged here is one value *de se* and not a multitude of values *de dicto* (Lewis, 2000 [1989a]: 75). *De se* values will not necessarily result in harmony:

"All might value alike, valuing *de se* the same properties and valuing *de dicto* the same propositions. Insofar as they succeed in desiring as they desire to desire, they will desire alike. But that does not ensure that they will agree in desire". (Lewis, 2000 [1989a]: 76)

Despite the disharmony that the pursuit of attitudes *de se* may bring, for Lewis (2000 [1989a]: 76), putative values *de se* are genuine, therefore the term must be built in our theoretical ethical framework. This explanation is useful, not only for its contribution to the Lewisian value theory scheme, but also for understanding better Lewis's method of analysis and systematization of folk moral psychology by using theoretical terms and modal notions borrowed from his modal metaphysical system.

3. Objection: Lewisian modal realism does not offer an alternative account of possible worlds, when compared with other modal realist accounts termed as 'ersatzism'.

Reply: For understanding Lewis's arguments against ersatzism, it is important to accept that Lewisian possible worlds (1986a: 84) are whole worlds and not parts of a greater world, because they are spatiotemporally and causally isolated from one another. Lewis (1986a: 78-80) argues that the causal isolation of worlds follows automatically under his counterfactual analysis of causation, by virtue of which he offers a good description of his concept of possible worlds – also apparent in the example of a possible world, where Aristotle is not a philosopher. Nevertheless, as we saw, the lack of spatiotemporal or causal relation between possible worlds, hence the lack of transworld causation, as well as of our standing in a world as individuals, does not justify assuming abstraction (Lewis, 1986a: 84).

In the example of transworld individuals, apart from the problem of causation between causally isolated worlds, Lewisian modal realism encounters objections by ersatzism, in terms of the acceptance of haecceitistic differences between worlds and individuals in worlds. Haecceitism, which is a modal thesis addressing ontological concerns about, amongst other things, the essentialism of properties, as well as qualitative and non-qualitative differences between objects and between worlds, supports concrete, yet ersatz, meaning abstract, worlds (Lewis, 1986a: 148). Lewis (1986a: 220-221) illustrates the haecceitistic position in reference to Kaplan's definition of "being extended in physical space and time". Conversely, anti-haecceitism does not accept transworld being for entities of distinct possible worlds. This view counteracts a rigid type of metaphysical determinism and its denial that the same thing exists according to many worlds, whilst also counter arguing the supervenience thesis (Lewis, 1986a: 221-222). For instance, in Lewis's view (1986a: 227), Kripke and himself, despite loosely accepting haecceitism, in fact they adopt slightly different positions: while Lewis rejects haecceitistic differences between worlds, so he is anti-haecceitist in Kaplan's sense (Lewis, 1986a: 221), Kripke does not accept such differences, adopting a more neutral position. Hence, Lewis (1986a: 70) objects to Kripke's claims of transworld spatio-temporal relations – for instance, in the example of telescoping viewing of spatio-temporally

connected worlds – while he is not against transworld comparisons.¹¹³

Lewis's argument concerning haecceitistic differences derives from his view about what corresponds to haecceitistic differences in his genuine modal realism; for Lewis (1986a: 159), these are "genuine differences in how things might be", because they consist in "differences in what properties things might have". In Lewisian modal realism (1986a: 159), a linguistic description of properties does not allow for a position, from which we can describe, for example, "completely any possibility in which there are extra natural properties, alien to actuality". Lewis (1986a: 162) argues that, if Ramsified, such natural properties are described incompletely. However, we have names for properties the ersatzist cannot name.¹¹⁴ Lewis (1986a: 161) calls them "missing properties" in our world and in the ersatz world, while he mentions fundamental properties alien to this world as examples of sparse properties of this-worldly, but also of other-worldly things, which problematize the construction of ersatz worlds (Lewis, 1986a: 60).

It is here that Lewis (1986a: 160) identifies a basic problem in ersatzism. Despite his and the ersatz world's acceptance that some properties rest beyond our thought and language (Lewis, 1986a: 179), for the modal realist, there are alien natural properties, or universals, which are only instantiated in other worlds (Lewis, 1986a: 167). Conversely, for the linguistic ersatz world, there are no other properties instantiated in other worlds, simply because there are no other worlds

¹¹³ Lewis (1983a: 20) explicitly argues against haecceitism, as a position examined in his counterpart theory, with respect to Kripke's discussion of cross-identification, as partly haecceitistic and in favour of essentialism, which, for Lewis, does not succeed in cross-identifying the subject by description. Lewis (1983a: 17-18) explains that, in settling questions of identity, indeed Kripke does not favour cross-identification through telescoping viewing: i.e. through a "Verneoscope", which will help us identify Nixon as an inhabitant of a world, where he is a politician winning the elections. Instead, he suggests that we consider a world, in which Nixon loses the elections, therefore, cross-identification occurs by stipulation. Kripkean stipulation of worlds and individuals is not a mental act of world-making, or identification-making, but rather a selection process:

"Out of all the worlds there are, we stipulate which ones we wish to consider. And in doing this, Kripke insists, it is perfectly right and proper to specify worlds in terms of cross-identification." (Lewis, 1983a: 18)

The Lewisian objections to Kripke's stipulation in terms of cross-identification (1983a: 20-21) concern, amongst other, the problem of selection when there are too few candidates, as well as the need to impose essentialistic constraints upon identity across worlds. Nevertheless, Lewis (1986a: 222) recognizes that the Kripkean specification, which is the specification of worlds by making reference to individuals does not enter any haecceitistic debates. The Lewisian philosophy of identity (1986a: 201) admits to essentialistic considerations with regard to intrinsic properties, such as being "essentially human", but not to extrinsic properties, which are accidental, such as how many dogs one owns in different possible worlds.

¹¹⁴ Using Carnapian terms, Lewis (1986a: 155) raises objections against linguistic ersatzism, by pointing at "a problem with consistency regarding world-making languages, when they speak of both local and global matters"; for instance, in the formulation of conditionals with local antecedents and global consequents. Hence, Lewis (1986a: 156) criticizes the linguistic ersatz world's view that the analysis of modality has to wait on an analysis of the global in terms of the local.

(Lewis, 1986a: 160); or, in the case of pictorial ersatzism, properties are instantiated as part of the ersatzer's abstract picture (Lewis, 1986a: 169). Lewis also comes to refine his philosophical problematization of the abstract and the concrete in terms of worlds; in saying that some pictorial ersatz "are literally abstractions", Lewis (1986a: 172-173) argues that the concrete world has a special ingredient, which is totally absent from all the ersatz worlds. The latter view complies with the extreme Lewisian modal realist view, in which things that are concrete parts of this world, have perfect duplicates that are concrete parts of other worlds. Such recognition, for Lewis (1986a: 82), settles the notion of concreteness, with regard to at least some possible individuals and some possible worlds.

Lewisian modal realism draws upon Lewis's ontological view (1986a: 163) that perfectly natural properties, which are intrinsic *ex officio*, can never differ between duplicates.¹¹⁵ Lewis (1986a: 221-235) further explains his objection to ersatz individuals and haecceitistic differences, based on his modal realist view of differences in representation *de re* of some individual in different worlds.¹¹⁶ For Lewis (1986a: 235), representation *de re* is done in worlds, not by worlds themselves in the canonical sense, but by the appropriate individual possibilities available within different worlds. Lewis (1986a: 235) explains that, "one world will provide at most one accessible possibility for a given individual". Introducing the notion of counterparts to account for duplication, Lewis also accounts for *possibilia* of individuals in variations of worlds. Seemingly attempting to replace transworld identity with counterparts, Lewis (1986a: 88) however argues that this is not the case.¹¹⁷ Lewis (1986a: 89) is clear that he takes counterparts together with their

¹¹⁵ This Lewisian view of intrinsic properties (1986a: 163) also supports Lewis's objection to combinatorialist ersatzism. Lewis (1986a: 90) credits combinatorialism with connecting imaginability with possibility; yet he objects to imaginability as a criterion of possibility:

"We can imagine the impossible, provided we do not imagine it in perfect detail and all at once. We cannot imagine the possible in perfect detail and all at once, not if it is at all complicated". Nonetheless, for Lewis (1986a: 90), we can take imaginative experiments as a means of informal reasoning from the recombination principle. Following the argument for abstraction, Berto and Plebani (2015: 174) argue that combinatorialists, such as Quine and Creswell, still take possible worlds as worlds of actuality: "Arrangements and rearrangements, or combinations and recombinations, are abstract things, so worlds still count as abstract".

¹¹⁶ The *de re* and *de dicto* representations or interpretations derive from medieval logic. A *de dicto* interpretation says that a proposition is necessary, while a *de re* interpretation says that a thing (*res*) has a property necessarily; see (Fitting and Mendelsohn, 1998: 86).

¹¹⁷ In lay terms, Lewis (1968: 115) defines counterparts as "men you would have been, had the world been otherwise". Using the notion of similarity, Lewis (1968: 114) explains that counterparts resemble an individual in context and content in important respects; in fact, they resemble that individual "more closely than do other things in their world", but they are not really that individual, "for each of them is in his own world, and only you are here in the actual world". The notion of similarity, or "sameness", as Lewis (1968: 115) argues, is here employed no more as "a literal identity than the

surroundings, therefore counterparts are world-bound, in this way also differentiating counterparts from transworld individuals, who require overlap of worlds (Lewis, 1986a: 216), but also duplicates, who are dependent on shared properties, whilst also being world-bound;¹¹⁸ as Lewis argues, "differently situated duplicates do not share all their properties". Duplicates might differ extrinsically in relation to their surroundings. Lewis (1986a: 89) mentions the example of "the molecules of a cat, which are perfectly natural properties, or defined intrinsic properties", but "might not be part of a cat in another world".¹¹⁹

In terms of *de re* representation of individuals and accounting for haecceitistic differences in his counterpart theory, Lewis (1986a: 235) raises several objections to different types of ersatzism and their degrees of congeniality to haecceitism. The Lewisian evaluation (1986a: 259-260) derives from the modal realist position that there is no overlap between worlds, counterpart relations are qualitative and thus representation *de re* is inconstant. Likewise, in haecceitism, there is no overlap between worlds, but counterpart relations are non-qualitative. Both in modal realism and in haecceitism, there is distinction of possibilities without qualitative difference of worlds; yet where the modal realist sees different possible individuals within the same world, the haecceitist sees different worlds (Lewis, 1986a: 243).

The above can be contextualized in the broader Lewisian critique of ersatzism. For Lewis (1986a: 150), modality must be taken as primitive for the ersatzer, because he or she "does not know the facts of modality" (Lewis, 1986a: 155). For the ersatzer, "what is true according to an ersatz world is what is true of this world"; and what is true is that the ersatz world is constructed "in a certain way from constituents of certain kinds" (Lewis, 1986a: 174). To account for genuine modality, which is not achieved in ersatzism,¹²⁰ in Lewis's modal realist view, one

sameness between you today and you tomorrow". As Lewis (1968: 115) explains, his counterpart theory does not follow other transworld theories, for instance by Hintikka (1962) and Kripke (1980), who propose interpretations of quantified modal logic, "on which one thing is allowed to be in several worlds". This distinction also justifies Lewis's argument that counterpart relation is not an equivalence relation.

¹¹⁸ Lewis (1986a: 205) also argues that whenever there is duplication, there is recurrence of universals.

¹¹⁹ Lewis (1993: 69) has argued against the conflation of intrinsic properties with relations; for instance, "the familiar property of being a swan into some unheard-of-swanning-at relation of things to worlds". Lewis (1993: 70) proves this by following a sentential expression with a superscripted expression denoting a world, which is defined in turn by a recursion.

¹²⁰ Lewis (1986a: 155) claims that the 'ersatzer' does not know "the facts of modality":

must take very few fundamental properties and relations as primitive. For Lewis (1986a: 155), this approach combines ontological benefits – “a safe and sane ontology” – with a bit of, instead of extensive, primitive modality. For the ersatz, either linguistic or pictorial, there is a contest between ‘genuine’ and ‘ersatz’ modal realism, which results in a choice between “unwelcome ontology and unwelcome primitive modality”. Lewis (1986a: 156) argues, however inconclusively, that ersatzers show a preference for the latter.

We saw that, for Lewis (1986a: 187), propositions are sets of worlds, while for the ersatzist they are simply sets of ersatz worlds.¹²¹ The Lewisian sets of worlds cannot be considered as abstractions, unless they are equivalence classes “only under thoroughly artificial equivalences” (Lewis, 1986a: 85). As Lewis (1986a: 85) argued: “I emphatically do not identify possible worlds in any way with respectable linguistic entities; I take them to be respectable entities in their own right”. Lewis’s proposal for the acceptance of a little primitive modality (1986a: 155), even if only reduced to extensional entities, accounting for genuine modality and safeguarding ontological concerns, has raised objections from his opponents.

Lewis (1986a: 174) also responds to those he terms as ‘magical ersatzers’, a modal realist strand in which he identifies several problems regarding modality and *de re* representation, primarily due to the lack of an account, on their part, of how representation works. For Lewis (1986a: 183), so-called magical ersatzism is this-worldly nondescript ‘ersatzism’: Plantinga and Stalnaker do not say how representing works (Lewis, 1986a: 141). In Lewis’s view (1986a: 143), their approach either misrepresents different possibilities as the same, or some possibilities as impossible, which simply gets the facts of modality wrong. Lewisian modal realism (1986a: 162) requires, not only for different possibilities not to be omitted, but also not to be conflated with one another. Lewis argues that where we ought to have many possibilities for an individual, for the ersatz, we have only one linguistic or pictorial “description of the filler of the role”; or we risk having an irrelevant multiplicity of ambiguous representations by magical ersatzism (Lewis, 1986a: 158).

“If the ersatz knew the facts of modality, he would know what to do. But he doesn’t and he can’t. What is he supposed to do – make them up out of whole cloth? That gets the facts wrong even if he guesses right! The only safe course is to resort to primitive modality”.

¹²¹ Lewis (1986a: 185) emphatically states that he is not opposed to “states of affairs, ways things might be, possibilities, propositions or structures”; but, for him, this means that he believes in entities that are suited to play the roles, which their names ascribe to them. His proposed entities are sets of worlds and are the same in every case: “Worlds as I understand them: us and all our surroundings, and other things like that” (Lewis, 1986a: 185).

4. Objection: The Lewisian modal realist thesis of the plurality of worlds does not contradict metaphysical actualism.

Reply: Lewis (1986a: 100-101) recognizes that metaphysical actualism is a substantive thesis about what there is; however, for him, actualism is a trivial analytic truth. According to Lewis, "to be necessarily (possibly) true is to be true at all (some) worlds, that is all (some) disconnected space-times" (Berto and Plebani, 2015: 170). In Lewisian worlds (1986a: 26), "a theory T defines a region in the space of possible worlds: namely the class of all T-worlds". This claim takes us to Lewis's thesis (1986a: 92) against actualism – "ours is one of many worlds" – which can be summarized in the position that there is no world alone that is absolutely actual.

Construed from the possibly raised objections:

(1) We cannot know with certainty that we are the select few, who live in an actual world, while others live in worlds that lack absolute actuality;

(2) It is a contingent matter which world is actual and a contingent matter varies from world to world;

Lewis (1986a: 93-94) argues that his modal realist view has certain implications, such as, as we saw, that the question of existence is not a genuine problem and that quantification can also be applied to non-actual things (Lewis, 1986a: 95). The latter may include properties, "taken as sets of their this- and other-worldly instances", yet also universal numbers and mathematical objects, which can be called actual, not by courtesy, but because they are parts of this world. For Lewis, this position also raises the problem of relativity, which he addresses by analyzing the indexical 'actual' and whether it is rigidified. For instance, comparing the terms 'now' and 'the present', Lewis (1986a: 94) suggests that while 'now' is rigidified by reference to the time of utterance, 'the present' is sometimes rigidified and sometimes not, since its reference is either unshifted, or may shift. Lewis (1986a: 94) understands his notion of 'actual' in possible worlds in the latter sense, as a cognate to the term 'present'. Likewise, properties are sets of this and otherworldly instances, yet only the properties whose instances are confined to the actual world are actual (Lewis, 1986a: 95). In the case of propositions, Lewis (1986a: 95) argues that, "we might call all propositions actual, distinguishing however between those which are true and those which are not actually true". In the above examples of temporal terms, one easily sees that Lewis uses the term actual as both a rigidified and a non-rigidified indexical.

5. Objection: Lewis's application of conditionals in his system of sphere semantics does not solve the problem of relativity.

Reply: The Lewisian system of spheres of possible worlds allows Lewis to work with a strict and a variably strict conditional, based on whether the system is universal, as the set of all possible worlds, or not. It consists of a set of worlds, or set of spheres U , around a world i , to which the outer modalities, outer necessity and outer possibility, pertain, being interpretable by means of accessibility of the outermost sphere, $U\$i$, to the world i (Lewis, 1973: 23). $\$$ is taken to be a (centered) system of spheres, $\$i$ a system of spheres centered on i , and the members of each $\$i$ are called spheres around i , iff certain conditions hold for each world i : such as centering, nesting, closure under unions, closure under intersections (Lewis, 1973: 14). Lewis (1973: 23) explains that if the system of spheres is universal, in which case $U\$i$ is the set of all possible worlds, the outer necessity and possibility will be ordinary logical necessity and possibility. If the system of spheres is not universal, then the outer modalities will be strict modalities: "probably stricter than anything familiar except the logical modalities themselves" (Lewis, 1973: 23). Lewisian spheres form a dynamic system, in which the counterfactual cannot be defined by the outer modalities, because the truth value of sentences changes in the system, when different spheres are added or deleted, yet without changing the outermost sphere around any world¹²².

Lewis (1973: 23) explains how the counterfactual conditional operates in his system of spheres. Taking $\supset$ as the entailment, or material implication, and $\Box$ and $\Box \rightarrow$ as the special modal operators, we have:

(1) the outer strict conditional if necessarily (ϕ entails ψ) implies the counterfactual if ϕ then necessarily ψ . This is formulated as:

$$\Box(\phi \supset \psi) \supset (\phi \Box \rightarrow \psi)$$

(1) holds.

On the premise (1):

(2) if the antecedent of (ϕ entails ψ) is true throughout every sphere around i , then it is true throughout the ϕ -permitting sphere around i , if there is such a sphere. Defining a function f which selects, for any sentence ϕ and world i , the set $f(\phi, i)$ of closest ϕ -worlds to i , let $f(\phi, i)$ be the set of ϕ -worlds belonging to every ϕ -

¹²² See figure 2 (Lewis, 1973b: 18). Here, worlds are taken as individuals in Lewisian terms, which means they are concrete and not abstract; see also Lewis's response to the conflation argument (1986a: 83).

permitting sphere in $\$i$, if there is any ϕ -permitting sphere in $\$i$, or the empty set otherwise¹²³. Taking that $f(\phi, i)$ is included in $[[\phi]]$ in the case there is some ϕ -permitting sphere in $\$i$, $f(\phi, i)$ is included in the intersection of $[[\phi]]$ and any ϕ -permitting sphere, it is sufficient to find a sphere S in $\$i$, which fulfills the following condition¹²⁴:

$$(I^*) f(\phi, i) \text{ df } = [[\phi]] \cap S$$

Taking a sphere S , a nonempty set I and based on a system of spheres $\$$ over i , let $[[]]$ to be a function mapping all sentences of the language onto the subsets of I and let i to be some particular world in I ¹²⁵, and taking condition (I^*) , (2) holds¹²⁶.

- (3) if the antecedent of $(\phi \text{ entails } \psi)$ is true throughout some ϕ -permitting sphere, but false in some larger ϕ -permitting sphere, then if ϕ then necessarily ψ is true and necessarily $(\phi \text{ entails } \psi)$ is false.

Taking into consideration the analysis for premise (2), (3) can be formulated as:

$$(\phi \Box \rightarrow \psi) \supset \Box(\phi \supset \psi);$$

hence (3) brings premise (1) false¹²⁷.

Using possible worlds in his semantics of counterfactual conditionals by means of his spheres, Lewis construes a deontic scheme partly out of his notion of comparative similarity of worlds. Lewisian counterfactuals are related to a kind of strict conditional based on comparative similarity of possible ϕ -worlds; a counterfactual: "if ϕ then it is necessary that ψ is true at a world i if and only if ψ holds at certain ϕ -worlds"; yet not all ϕ -worlds matter¹²⁸ (Lewis, 1973: 8). However, as shown above, Lewis (1973: 4) argues that the counterfactual cannot be any strict conditional.

Lewis (1973: 92-93) considers this view significant for his analysis of the vagueness of counterfactual conditionals in possible worlds, despite recognizing that comparative overall similarity is vague, while it may consist of "innumerable similarities and differences in innumerable respects of comparison, balanced against each other according to the relative importances we attach to those respects of comparison". Lewis (1973: 1) takes counterfactuals themselves to be vague,

¹²³ See (Lewis, 1973: 58).

¹²⁴ Ibid., 58-59.

¹²⁵ Ibid., 118.

¹²⁶ Ibid., 23.

¹²⁷ Ibid., 23.

¹²⁸ See figure 3 (Lewis, 1973: 17).

especially if they are true conditionals based on similarity (Lewis, 1973: 12)¹²⁹, by ascertaining that truth conditions for counterfactuals are only fixed within rough limits, as well as vary with contextual changes and with regard to the relativity of degrees and respects of comparison (Lewis, 1973: 92). Nevertheless, the limited vagueness of counterfactuals accounts for the limited vagueness of similarity: "they are vague in a coordinated way" (Lewis, 1973: 92-94). Lewis (1973: 92) explains that, for giving precise conditions of truth relative to any given value of a delineating parameter, which consists in Lewis's precisification (2000 [1989a]: 86), we need to give values to the delineating parameters for solving the problem of relativity.¹³⁰ In counterfactuals, the delineating parameter for defining the vagueness of counterfactuals is the comparative similarity relation itself, which indicates the system of spheres, the comparative similarity system, as well as the selection function; or, as Lewis (1973: 93) mentions, "whatever other entity we use to carry information about the comparative similarity of worlds".

Furthermore, Lewis (1973: 100) constructs deontic conditionals, as versions of the conditional and unconditional modal operators of obligation and permission. Depending on whether there is an innermost nonempty sphere around each world *i*, which contains the best of all worlds from the standpoint of *i*, these deontic conditionals may be treated as ordinary necessity and possibility operators, which can be interpreted by assigning the best worlds to *i* as its sphere of accessibility. If there are no best worlds from the standpoint of *i*, there is instead an infinite ascent to better and better worlds; at which the truth value of the necessity deontic operator is strictly conditional upon the variable of the antecedent holding throughout all sufficiently good worlds and the truth value of the possibility deontic operator is strictly conditional upon there continuing to be some worlds no matter how high one ascends (Lewis, 1973: 101).

¹²⁹ Lewis (1973b: 18) proves that the sequences of true counterfactuals and their true negations led him to reject the theory of the counterfactual as a constantly strict conditional based on similarity.

¹³⁰ As it is the case with equivocal analyticity, which I discussed in **subsection 3.1**.

Appendix II

Glossary¹³¹

Absolutism: view abiding with conceptions of the absolute as a thing, or an entity, possessing maximal ontological status.

Moral absolutism: view that an act is right or wrong, good or evil, regardless of the context of the act (i.e. cases of homicide, the execution of an innocent person); view that actions are inherently moral, or immoral.

Accessibility relation: modal term used in possible worlds semantics (PWS) for a relation on a set of possible worlds, which restricts quantification over possible worlds.

Actualism: the modal metaphysical notion of distinguishing between actual and possible, or merely possible, objects, or worlds; the view in PWS that there is only one world that is absolutely actual.

Alternative: Lewisian term for a class of counterparts.

doxastic: Lewisian term for a class of an individual's alternatives in respect to a belief system in possible worlds.

epistemic: Lewisian term for a class of an individual's alternatives in respect to an epistemic system in possible worlds.

Analyticity: term for the philosophical problem pertaining to the distinction between the analytic and the synthetic; Carnapian-Quinean term, also used by Lewis, regarding the truth value of logical, mathematical and set-theoretical statements.

Belief-desire: the moral psychological motivation of an agent's actions combining belief and desire.

Besire: attitude of valuing that is both a belief and a desire; abbreviation for belief-desire.

Branching: term used in counterfactuals using tense and modal logic to formulate temporal conceptions; Lewisian term used in PWS to describe worlds, which share one initial spatio-temporal segment.

Category: term for genera of entities.

¹³¹ This is a cursory list of philosophical and other technical terms included in this thesis, which comprise of standard terms used in philosophy of language, modal metaphysics and canonical ethics.

Causation: causal, or cause and effect, relation between two elements, actions, or events.

Compatibilism v. incompatibilism: view that free will can co-exist with determinism v. view that free will is not compatible with determinism.

Conditional: logic term for hypotheses and propositions formulated by an antecedent and a consequent.

Consequentialism v. non-consequentialism: view that the evaluation of an act, or course of conduct, must be based on its consequences, or outcome v. view that a course of conduct is not to be evaluated solely on its outcome.

moral: judging an agent's action based solely on its consequences, rather than on the moral quality of the action.

Conventionalism: view that the actions of agents conform to regularities, relative to conventions, such as rules, norms, customs, practices, institutions and social contracts.

Counterfactual: conditional hypothesis, or proposition, based on a false, or invalid antecedent.

Counterparthood: Lewisian term used in PWS for cross-world identity based on a weak relation of similarity.

Deontology (deontic theories): domain of moral theories concerned with obligation and permission regarding an agent's actions; what one ought to do.

Determinism v. non-determinism: view that events, or actions, are causally determined by antecedent events, actions, or other factors v. view of the lack of causal necessitation between events, or actions.

soft-: view that, although there is some causal determination of events, or actions, an agent's actions are free, when the agent is said to have acted freely.

Deterrence: judicial term for the prevention of an injurious act.

Divergence: Lewisian term used in PWS to describe, instead of one world sharing a common segment, two worlds with duplicate initial segments; scientific term from the many-worlds interpretation (MWI) in quantum mechanics.

miracle: Lewisian term used in Lewis's theory of causation to define an occurrence (i.e. act, event, natural phenomenon) causing the 'divergence' of actions, events, or phenomena; on the Lewisian account of counterfactuals, the infamous example of Nixon's button-pressing world is interpreted in respect to the claim that the actual world is more similar to a button-pressing world diverging

from the actual world by means of a 'miracle' than the actual world converging with a button-pressing world by means of a 'miracle'.

Error theory: cognitivist moral philosophical theory, supporting the view that ascriptions of moral qualities to acts, events and objects are systematically and uniformly false.

Equilibria: game theoretical term pertaining to solution concepts for games amongst instrumentally rational agents for optimizing each agent's strategy in response to other agents' strategies; traditional utilitarian game theoretical view.

coordination: game theoretical term for expected behavior amongst agents.

Ersatzism: Lewisian term for modal realist metaphysical theories, which do not abide with the Lewisian view of genuine, or extreme, modal realism in possible worlds; the more generic term of abstractionism has also been used for ersatz theories; representationalist theories.

linguistic: Lewisian modal term for the strand of modal realism abiding to the view that linguistic descriptions stand for possible worlds.

pictorial: Lewisian modal term for the strand of modal realism abiding to the view that pictorial (i.e. isomorphic, a model-theoretic term standing for a relation between distinct objects) descriptions stand for possible worlds.

magical: Lewisian modal term for the strand of modal realism not presupposing an account of how representation works in possible worlds, by reliance on the structure, mereological, or set-theoretic, of the world.

Essentialism: view pertaining to the essentialist nature of different species.

Externalism: linguistic theory for developing models of the structural properties of language, such as speech sounds, words, and various linguistic elements.

Identity: sameness; a notion relating to a sense of personhood.

transworld: modal realist term for the identity of an individual determined by the conditions, or metaphysical principles, under which an individual relates to one world, and those under which the individual relates to a second, or *n*-world.

world-bound: Lewisian term for the identity of an individual, who is bound to the only one world, inhabited wholly by the individual.

Fictionalism: philosophical discourse, in which statements put forth need to be taken as analogous to fictional statements, rather than as literal truths.

modal: approach to statements and propositions about possible worlds.

Folk psychology: term used by philosophers and psychologists to describe our ordinary mentalistic notions.

Free will: a significant control over one's actions.

Haecceitism v. anti-haecceitism: the modal thesis supporting the view of the qualitative, without the non-qualitative, differentiation of a world, or of the non-determination of identity due to lack of essential, or necessary properties v. the modal thesis that there is necessarily qualitative with non-qualitative differentiation of a world, or of the determination of identity by essential, or necessary properties.

Humeanism v. anti-Humeanism: view that there is no prioritization of belief over desire in the motivation of an agent's intentional actions v. view that primarily beliefs, or beliefs with their corresponding aligned desires, motivate an agent's intentional actions.

Indeterminacy: vagueness; fluidity of meaning.

semantic: the problem in analytic philosophy of the non-fixity, or of the precision of meaning.

Induction: the problem of drawing upon known instances to make general, or probabilistic claims.

inductive reasoning: type of probabilistic reasoning based on general principles of induction.

Intentionality: term from the realm of the philosophy of mind used to address ontological and metaphysical questions about mental states.

Materialism: strand of philosophy pertaining to explanations relating to matter, as understood and examined in physics and in the natural sciences.

Metaethics: area of moral philosophy concerned with questions about a question of normative ethics, but also with broader questions including semantics, methodology, epistemology, or metaphysical and psychological questions.

Metaphilosophy: area of study concerned with questions about the nature of philosophy.

Metaphysics: area of philosophy dealing with metaphysical problems, such as, amongst others, the problem of being, of free will, and of the mental and the physical.

modal: strand of contemporary metaphysics dealing with modal problems of necessity, contingency and possibility.

Moral psychology: a combination of ethics and psychology.

folk: a combination of ethics and psychology of the common person.

Moral rule: rule, or principle, in moral philosophy.

necessary: modal term for the necessitation of a moral rule, or principle.

conditional: modal term for the conditionality of a moral rule, or principle.

Modal: logic operator specifying the mode, in which a sentence, or proposition, is said to be true.

box: modal operator for 'it is necessary that'.

diamond: modal operator for 'it is possible that'.

Modality: notion for statements, or propositions, concerned with problems of necessity, contingency and possibility (i.e. metaphysical, epistemological, nomic).

de dicto: the modality of propositions, concerned with modal properties "said of", or ascribed to, an object.

de re: the modality of objects, concerned with modal properties that an object itself has per se.

de se: Lewisian modal term for the egocentric modality of propositions and objects, concerned with (self-)ascriptions of propositions and properties.

deontic: notion for modal statements where the modal 'necessary' is read as 'it is obligatory' and the modal 'possibly' is read as 'it is permissible'.

doxastic: notion for modal statements where the modal 'necessary' is read as 'it is believed'.

epistemic: notion for modal statements where the modal 'necessary' is read as 'it is known'.

Naturalism: philosophical notion allying philosophy with science, under ontological and methodological naturalistic doctrines.

Nomological: statements pertaining to laws (i.e. natural, scientific, societal).

Ontology: the area of metaphysics concerned with questions of being.

Perdurantism: the metaphysical view that objects, or individuals, exist through temporal, or modal parts, which are never wholly present at more than one time, or world, across time or possible worlds.

Property: ontological term for a specific quality, or feature, of an object pertaining to the metaphysical problem of individuation.

intrinsic: property, which is entirely about a thing itself.

essential: Lewisian term for a property of all counterparts of an individual.

extrinsic: property, which a thing has, in virtue of its environment.

Positivism: philosophical system pertaining to the view of scientific, logical, or mathematical verification of the authenticity of knowledge.

logical: strand of analytic philosophy pertaining to the view that the only sources of knowledge derive from a combination of the *a priori* analytic and the *a posteriori* synthetic, specifically logical reasoning and empirical science; the rejection of *a priori* synthetic.

Possible worlds: modal notion and term standing for a hypothesis used in analytic metaphysics.

Possibilism: metaphysical view pertaining to distinctions between being and existence.

classical: metaphysical view distinguishing between being and existence, or actuality.

modal realist: metaphysical view not recognizing any distinction between being and existence.

Possibilia: things that could have existed, but do not exist, in classical possibilist terms; things that could have existed, do exist, in Lewisian modal realist terms.

Projectivism: metaethical view that evaluative properties are subjective projections caused by sentiments, such as emotions and moral psychological attitudes.

Quantification: the philosophical logical method of analysis using quantifier expressions (i.e. 'all', 'some', 'many', 'most' etc.).

existential: quantification applied in the realm of classical ontology.

possibilist: quantification applied in modal realist ontology.

Radical interpretation: the problem in analytic philosophy of determining the truth value and meaning of speech amongst speakers of the same language, taking into consideration psychological and behavioristic factors.

translation: the problem of radical interpretation from the viewpoint of an interpreter speaking a different language than the speakers whose speech is interpreted, but also possessing prior linguistic knowledge of the actual language, or of some related language.

Rationality: quality of being rational, or agreeable, to reason; acceptance of reason as the only source of knowledge.

instrumental: a sort of practical reasoning concerning the efficient use of appropriate means for achieving certain outcomes.

Realism v. anti-realism: view that objects and possible worlds have a mind independent ontological existence v. view that objects and possible worlds have a mind dependent ontological existence; view supporting actualism.

irrealism: metaethical term used by Lewis to define his theory of value as dispositional.

modal: metaphysical and ethical term used in PWS.

quasi-: metaethical term supporting the view that moral statements are not propositional and are projective of moral attitudes.

Reductionism v. non-reductionism: analysis and description in terms of simple or fundamental constituents for providing sufficient explanation of a phenomenon v. the opposite of reductionism.

Relativism: philosophical doctrine supporting conventionalist and contextualist interpretations of truth and falsity, right and wrong, or standards of reasoning.

metaphysical relativism v. metaphysical realism: context-dependent or mind-dependent view of the existence of objects, or metaphysical truths v. mind-independent existence of objects, or metaphysical truths.

Selection function: conditional logic term used for quantification at possible worlds.

Self-determination: characteristic of personhood pertaining to an acting subject's sense of efficacy on their part; compatibilist and incompatibilist moral philosophical term.

Semantics: area of linguistics and the philosophy of language concerned with the study of meaning.

sphere: Lewisian term for the PWS semantics used in his philosophical system of spheres.

population: Lewisian term for semantics in a population.

possible worlds (PWS): modal term for the semantics introducing the 'possible worlds' notion in modal logic and used in the possible worlds hypothesis to make claims about the conditions of the truth value of sentences at possible worlds.

Similarity: Lewisian term for defining similarities amongst possible worlds and possible world-bound individuals in intrinsic respects, also implying the context-dependency of counterfactuals.

Subjectivism: moral philosophical view of the lack of objective moral truths.

Supervenience: notion consistent with the view that an entity, or a set of properties, supervenes upon another entity, or set of properties, if there cannot be any difference to the one without being a difference to the other.

psycho-physical: notion consistent with the view that there cannot be any difference of the mental without any difference of the physical.

Teleology: reason, or explanation, based on the presupposition of an end, purpose, or goal of a phenomenon or action.

Theodicy: theistic view of the intervention by an omnipotent and omniscient being, a deity, either for the prevention of evil acts, or for the restoration of justice.

Tolerance: moral principle advocating the conditional acceptance of, or non-interference with, beliefs, actions, and practices, towards which one has an evaluative objection.

Universalism: view of the existence of universal and objective truths.

Utilitarianism: strand of normative ethics concerned with maximizing expected utility, or good, based on the premise of agent-neutrality; a consequentialist theory.

act-: an action's moral status seen as a function of its utility.

extreme, or radical: applying utilitarian principles to everyone's actions.

rule-: an action's moral status determined by moral rules and secondary principles.

Vagueness: semantic ambiguity, or lack of semantic precision.

Value pluralism: view pertaining to the plurality of the things in question, such as moral or political values, concepts, and viewpoints.

Value theory: term pertaining to all areas of philosophy that encompass some evaluative view.

dispositional: moral philosophical, as well as social and political philosophical, term, supporting the view that evaluative attitudes are determined by what an agent is disposed to value.

Virtue ethics (*aretai*c theories): domain of moral theories concerned with the virtues of an agent; what kind of person one is, or ought to be.

Appendix III

Index of Lewis's symbolic notation¹³²

- : 'it is necessary that'
- ◇: 'it is possible that'
- iff: 'if and only if'
- >: entailment, probability revision conditional
- ⊃: implication, entailment, truth functional connective
- : necessary entailment, special modal operator
- A, Φ, ψ, P, Q, s: proposition, or statement, or formula
- σ: performed act
- P: probability function
- V: value function
- C: credence function
- V, U: variant conditionals
- φ-world: world at which proposition φ holds
 - P-, Q-, R-: worlds at which propositions P, or Q, or R hold respectively, where R also denotes a binary relation of accessibility over worlds; doxastically accessible worlds
- φ-permitting sphere: deontic sphere of accessibility where proposition of permissibility φ holds
- i, w, W: a world
- I: a set of worlds
- S: a sphere around a world, which consists of a set of worlds
- Si: a sphere of accessibility to world i, if certain conditions hold
- \$: a centered system of spheres
- \$i: a centered system of spheres on i
- {i}: a sphere in which world i is centered
- U: set of spheres
- U\$i: the outermost sphere of accessibility to world i
- O: obligation deontic operator
- A, B: sentences
 - O(A/B): deontic sentence
- L1, or L2: languages

¹³² This list is not a comprehensive list of the notation used by Lewis, but only includes notation by Lewis, which appears as it is referenced in this thesis.