



UNIVERSITÀ DEGLI STUDI DEL PIEMONTE ORIENTALE
“AMEDEO AVOGADRO”

DIPARTIMENTO DI STUDI UMANISTICI

Corso di Dottorato di Ricerca in Filosofia

Ciclo XXXVII

**“ON RADICAL MORAL CHANGE. MORAL PARADIGM SHIFTS
AND THE POSSIBILITY OF PROGRESSIVE VIRTUE ETHICS”**

SSD M-FIL/03 Filosofia Morale

Tutor

Prof.ssa Maria Silvia Vaccarezza

Dottoranda

Matilde Liberti

Coordinatore

Prof. Luca Fonnesu

Al nonno Sergio e al maestro Romano

Table of Contents

Acknowledgements p. 6

Introduction p. 7

Chapter I. Radical Moral Change p. 11

1. Moral Conversion p. 12

1.1 Epiphanic Moral Conversions p. 13

1.2 Moral or Psychological? p. 17

1.3 The ‘Moral’ of Moral Conversion p. 23

1.4 Moral and Religious Conversion p. 31

2. Ladders of Moral Transformation p. 40

2.1 Transformation as ascension p. 40

2.1.1 The Symposium p. 41

2.1.2 Ascension and Contemplation in Wittgenstein’s TLP p. 44

2.1.3 Moral Transformation in Murdoch p. 52

2.2 Transformative activities p. 61

3. Moral Progress p. 70

3.1 Individual Moral Progress p. 71

3.2 Wittgenstein on the illusion of progress p. 74

3.3 Moral Progress and Moral Revolutions p. 78

4. Conclusion to Chapter I p. 83

Chapter II. Moral Paradigm Shifts p. 96

1. Conceptual Framework p. 97

1.1 Active and Passive p. 98

1.2 Abrupt and Gradual p. 99

1.3 Peripheral and Core p. 101

1.4 Triplets p. 104

2. Moral Paradigm Shifts p. 110

2.1 Scientific Paradigms and Moral Paradigms p. 111

2.2 Scientific Paradigm Shifts and Moral Paradigm Shifts p. 114

2.2.1 Passive p. 116

2.2.2 Abrupt p. 118

2.2.3 Core p. 120

3. Looking for Moral Paradigm Shifts p. 123

3.1 Moral Conversion p. 123

3.2 Moral Transformation p. 126

3.3 Moral Progress p. 130

4. Conclusion to Chapter II p. 132

Chapter III. Virtue Ethics and Moral Paradigm Shifts p. 139

1. Moral Change in Virtue Ethics p. 143

1.1 Neo-Aristotelian Moral Development p. 146

1.2 The need for Moral Paradigm Shifts p. 149

1.3 The problem of Conservatism p. 153

2. The (in)flexibility of Virtues p. 155

2.1 Virtue Ethics and Tradition p. 156

2.2 Foot's Naturalism p. 162

2.3 Naturalism of second nature and metaphysics: Swanton's alternative p. 167

2.4 Ethical Naturalism without Metaethical Naturalism	p. 177
3. A Progressive Moral Outlook	p. 181
3.1 Moral Development and Epistemic Development	p. 183
3.2 Epistemic Virtues for a Progressive Moral Outlook	p. 186
3.3 Dynamic Flourishing	p. 192
3.4 Two Examples	p. 194
4. Conclusion to Chapter III	p. 204
Conclusion	p. 215

Acknowledgements

I sincerely thank my supervisor Prof.ssa Maria Silvia Vaccarezza for the constant and incredibly valuable support that she has always granted me throughout these research years made of frantic drafting, successful and unsuccessful submissions, acute and trivial intuitions, long meetings and, Aristotelically, just the right amount of desperation. I also wish to thank the academic staff of the FINO PhD program that has read and commented on my drafts during the WIP seminars and offered support every time it was needed.

I also wish to express my gratitude to my visiting supervisors, prof. Hanno Sauer (Utrecht University) and prof. Uriah Kriegel (Rice University), for their generosity in providing feedback on my drafts and for the invaluable gift they have offered me when they accepted to host me as a PhD visitor.

The best part of doing philosophy for work is that you are surrounded by colleagues who truly wish to exchange ideas and engage with newborn intuitions. I am extremely grateful to all my colleagues from the FINO PhD Ethical-Political curriculum who have helped this work take shape, as well as to my Genoa office colleagues who have offered me the opportunity to ping-pong ideas around the room for hours. This is truly the beauty of this all.

A sincere “thank you” goes also to all the non-official philosophers of my life that have supported me every day for the past three years (and more). In particular, I am grateful to my family: they are always there for me, and they truly believe that I can do this, even when I myself doubt it. A huge thank you also goes to my wonderful circle of friends: when I am most frustrated they make me feel loved and appreciated independently of anything else. This type of love is enough: the rest is just topping.

Introduction

“People don’t change.”

We see people around us so deeply embedded in their culture that, if we share the same culture, we can even attempt to make predictions about what they would say, think and vote for. We look at ourselves and realize that we are probably just as predictable as anyone else; we notice that we always make the same old mistakes – “why don’t I ever learn?”, “why can’t I be different?”, “why can’t she see that what she’s doing is hurting everyone else?”. Faced with these patterns, we end up thinking that no, “people don’t change.” Maybe in principle they can, but for some reasons they *don’t*.

Yet, narratives of all sorts – novels,¹ TV shows,² movies³ – are packed with stories of change. Not to mention that, when it comes to culturally Christian societies, we can see that the Bible offers a wide range of such stories,⁴ and even those who do not read the Bible are acquainted with these patterns of change, at least very generally. So, if there are so many such stories that even encompass cultural specificities, maybe people do change every now and then.

The first thing to notice is that “change” here has a radical flavour to it. Of course people do undergo change, in the sense that, physiologically, kids develop into teens and then into adults; hair grows, hairstyles changes, the tone of voice becomes deeper or more acute; we change tastes, preferences, sometimes even ideas about trivial matters, such as whether one would be interested in watching one particular movie rather than

¹ Classical examples are Jean Valjean in *Lés Miserables*, or Scrooge in *A Christmas Carol*.

² My favorite is Zuko’s story in the Nickelodeon Production Series *Avatar, the Last Airbender* (2005 – 2008).

³ See for example 1999 Tony Kaye’s movie *American History X*.

⁴ Take King David’s redemption arc in the first and second books of Samuel and Psalm 51, or Matthew’s call in, e.g., Luke 5, 27-32.

another. So yes, people do change something – physical aspect, ideas – in some respect. However, the narratives of change just highlighted are of a different kind: they involve a change that is *radical*, where radical is to be intended in the Latin sense: that is, it is a change in something that is rooted, something core, something that was so centrally valuable that it seemed impossible that it could ever be changed. In this dissertation I am interested in changes that are radical in this sense. Moreover, I am interested in changes that are both radical and morally significant:⁵ that is, that involve a fundamental change in one's action-guiding values, principles and ideals.

The argumentative structure of the dissertation can be summarized as follows:

(i) If Radical Moral Change (RMC) occurs, then it has morally significant implications for both agents and moral theories.

(ii) RMC occurs (Chapter I).

(iii) RMC has morally significant implications for both agents (Chapters I, II) and moral theories (Chapter III).

Chapter I unfolds around the following question: which forms can radical moral change take? There, I explore three main conceptualizations of radical moral change: moral conversion (section 1), moral transformation (section 2) and moral progress (section 3), and analyse possible common elements, as well as points of departure. I conclude that there are some common parameters according to which different forms of radical moral

⁵ One could wonder whether there is such a thing as a radical change that is morally *insignificant*. Think of changes that are radical for the individual, yet morally equivalent, e.g. I wished to become a parent, but I decide to become a nun instead. They are both morally acceptable options, but radically different for my life course. A change in this sense would be radical, but not morally significant, since the moral acceptability of the options is not significantly different. More on this in Chapter II.

change take shape. One parameter is that of abruptness/gradualness; for example, “epiphanic” moral conversions are characterized by an element of abruptness, and the same goes for “revolutionary” moral progress, while moral transformations and some forms of moral conversion unfold as gradual, step-by-step processes. Another parameter is whether the change was expected or unexpected: moral conversions are generally unexpected, “received” phenomena, while moral transformation and moral progress are generally planned processes, or at least expected.

In Chapter II I provide a possible classification of radical moral change that revolves around three parameters: (i) whether the change is actively pursued or passively received, (ii) whether the change is experienced as gradual or abrupt, and (iii) whether the change is core or peripheral. I then isolate one particular form of radical moral change that I believe worth of attention, which is that of changes that are passively received, experienced as abrupt and that take place at the level of core commitments and values. These changes is what I call “Moral Paradigm Shifts.”

In Chapter III I advocate for a type of neo-Aristotelian Virtue Ethics that can account for the possibility for people to undergo Moral Paradigm Shifts. The necessity of this move is due to the realization that, as of now, Virtue Ethics as a family of theories seems particularly inhospitable for radical moral changes that are progressive, to the extent to which they are disruptive of previous moral paradigms. I will argue that it is both possible and necessary for Virtue Ethics, specifically in its neo-Aristotelian declination, to adjust in order to make room for a type of moral change that is not only status quo reinforcing – as it is for standard moral development – but also potentially status quo disruptive – as it is for moral paradigm shifts.

This journey will end with a gaze to the structural level of radical moral change. Individuals will be the main focus of this dissertation; but the call for people to develop

a progressive moral outlook that I develop in Chapter III requires the further step of considering the structural elements that come to light when one begins to look at the moral world through a more accurate paradigm. This, however, will be the focus for further research.

Chapter I. Radical Moral Change

This chapter has the following purposes:

1. Provide an overview of how radical moral change has been considered in philosophical literature;
2. Isolate recurrent parameters for conceptual analysis.

In particular, we will see that all conceptualizations of radical moral change provide different ways to understand the “radical” element: as a change in “self”, as a distancing from the “self”, and as a discovery that calls for reconfiguration.

In philosophical literature radical moral change has mainly been conceptualized in three ways: as “moral conversion”, as “moral transformation” and as “moral progress”. The aim of this chapter is to offer a critical review of these terms. We shall see that comparing these concepts leads to murky grounds. Are they all just the same thing after all? Why then call them with three different names? We will see that, although they all describe similar phenomena, there are noteworthy differences. These forms of change can be more or less radical, more or less linear, and present more or less similar phenomenology.

In (1) I will consider “moral conversion,” focusing in particular on what it is that makes it “moral” and what it is that makes it a “conversion”. In (2) I will explore two ways to make sense of “moral transformation:” the “ascension” analogy, which sees moral transformation as the process of ascension towards something that is beyond ourselves, and the “activity” reading of transformation, according to which one morally transforms through practices. In (3) I will consider “moral progress” in its individual, or agent-level

dimension. I will conclude by introducing a particular form of moral progress, which is that of “moral revolution,” and propose that something akin to societal-level moral revolutions takes place also within individuals; this is the most radical form of moral progress.

1. Moral Conversion

“Conversion” comes from the Latin word *con-vertere*, which can be translated as “turning (in)to”. Combined materials convert into other, different materials, e.g. coal, water and air convert into nylon. In a similar fashion we employ this word to indicate instances in which people seem to turn into someone else; or in which people, in looking for guidance of some sort, turn to something that is completely different from what used to guide them before. In this latter sense, “conversion” is commonly associated with religion, in its describing the phenomenon of changing one’s religious affiliation and, consequently, changing also fundamental aspects of one’s self. Following this line of thought, a moral conversion can preliminary be understood as a point of turn into a different moral agent, or a moment after which one changes one’s moral guidance.

As we shall see, moral and religious conversion have a lot in common, but it is important to – at least – attempt to keep them separate. To begin with, religious conversion is likely to imply also a conversion of values and ideals; thus, a religious conversion is very likely to imply a moral conversion. However, moral conversion does not necessarily involve religious conversion; that is, one can undergo a radical change in values and ideals without necessarily commit to any religious beliefs or affiliations. Thus, it makes sense to consider moral conversion as a phenomenon in and of itself, independently of any religious connotation. In what follows I will consider the most

recent account of “epiphanic moral conversions,” which is that of Kristjánsson (2018; 1.1), and then I will expand on the following underlying narratives that help us understand what it is that we have at hand: Dees (1996) and MacLaughlin’s (2008) analysis of the “moral” aspect of moral conversion (1.2, 1.3), and Griffin (1980) and James’ (1902/1958) analysis of the “conversion” aspect (1.4).

1.1 Epiphanic Moral Conversions

Recently, a comprehensive analysis of instances of radical moral change has been provided by Kristjánsson, who calls them “epiphanic moral conversions” (2018). He defines epiphanic moral conversions as phenomena in which “a coincidental event can create cognitive dissonance in people, strong enough to motivate them to change their overall moral outlooks in abrupt and radical ways, by envisaging horizons of moral goodness that were previously closed to them.” (Kristjánsson 2018: 17). He then goes on to isolate what it is that makes them peculiar phenomena:

1. “[...] epiphanic [moral] conversions constitute abrupt, swift, or even catastrophic turning points (cf. MacLaughlin, 2008 , chap. 1)” (2018: 18). Kristjánsson stresses that even if radical change can happen gradually, epiphanic conversions are characterized by “quantum psychological changes” (Bien 2004) or “Aha moments” (Irvine 2015).⁶

2. “Epiphanic” can be understood as referring to “awakenings (to purported new truths) that are not only abrupt in a temporal sense, but also spontaneous and dramatic,

⁶ As Kristjánsson notes, quoting Irvine, “[t]he relative swiftness of the actual turning point, however, does not exclude a period of gradual moral preparation that may be noticed when the conversion experience is viewed in hindsight (Irvine, 2015 , chap. 4)” (2018: 18).

involving radical reconfigurations of mental structures [...]” (Kristjánsson 2018: 18). Even if the agent undergoing the epiphany cannot herself make sense of what exactly it is that is happening to her, the experience is one of “sudden “irradiations” disclosing meanings that were previously “occluded” (Cottingham, 2012)” (ibid.).

3. “[T]he events eliciting the conversion will normally be unplanned and fortuitous, or what Irvine (2015 , p. 17) calls “unbidden.”” (ibid.). However, Kristjánsson adds, that of being unplanned is not a necessary feature of epiphanic conversions, since agents may willingly facilitate their self-change out of perceived need to do so, or they can be facilitated to undergo such change by others, e.g. educators (see Schinkel 2016). “Nevertheless, the results of any such self- induced or other-induced interventions will always retain some measure of uncertainty and fortuitousness, as epiphanies rely on spontaneous reactions, and spontaneity can, by definition, never be fully planned” (Dees 1996 , p. 533)” (2018: 18).

4. “[E]piphanic moral conversions are strongly emotionally laden. This can mean either that they are triggered by intense emotions, such as awe or elevation [...], or elicit fierce emotional responses during and after the conversion, or both.” (2018: 19).

5. “[T]hose conversions involve radical self-change” (2018: 19), where what “self-change” amounts to depends on one’s theory of the self. As a realist about the self, Kristjánsson explains this point as meaning that those conversions involve “an essential change of the deep, underlying self, perhaps constituted by emotions of which the agent is unaware” (ibid.). Sceptics about the existence of an underlying self may understand this as involving beliefs and attitudes about ones’ self, which does not commit one to metaphysical claims about the self. Obviously, there is a sense in which the agent remains the same self in a metaphysical sense (“we are not talking about Kafkaesque

change of a Fregor Samsa into a gigantic bug”, *ibid.*), but maybe not the same self when it comes to identity/personhood.

6. “[T]he change of heart triggered by the moral conversion must be towards moral improvement or reform.” (2018: 19), both internally and externally. That is, it is necessary that the radical change constitutes better outcomes both for the person’s well-being and when it comes to her attitudes towards the world.

Thus, epiphanic moral conversions are characterized by abrupt turning points that involve radical reconfigurations of mental structures, that are exogenous to the extent to which they cannot be entirely self-induced, emotionally laden and directed at better outcomes. Crucially, they involve radical self-change. Kristjánsson draws epiphanic moral conversion’s six characteristics from previous discussions on the topic of moral conversion, both from a philosophical and psychological perspective. In philosophy, the reference points are Richard Dees’ paper ‘Moral Conversion’(1996) and Alfredo MacLaughlin’s PhD Dissertation ‘Narratives of Hope’ (2008). When it comes to psychology, Kristjánsson notes the fact that Aristotelian moral psychology does not seem to offer much room for the possibility of moral conversion;⁷ similarly, scholars

⁷ Yu 2007 rightly points out that Aristotle does not have the patience to convert wicked adults. Aristotle: “The bad man, if he is being brought into a better way of life and thought, may make some advance, however slight, and if he should once improve, even ever so little, it is plain that he might change completely, or at any rate make very great progress; for a man becomes more and more easily moved to virtue, however small the improvement was at first. It is, therefore, natural to suppose that he will make yet greater progress than he has made in the past, and as this progress goes on, it will change him completely and establish him in the contrary state, provided he is not hindered by lack of time” [13a22–31]. But Kristjánsson is not convinced that this is enough to dismiss this possibility (2018: 23).

have read Kohlberg's 6-stage theory as pretty deterministic,⁸ while Platonic moral psychology seems to be more hospitable for the possibility of personal transformations of this kind.⁹ In this regard, we will come back to the alleged rigidity of Aristotelian approaches to moral change in Chapter III. Kristjánsson then draws a parallel with the phenomenology of religious conversion: he considers William James' *Varieties*, and stresses that, according to James, not only these kinds of radical changes are possible, but they can happen both exogenously or endogenously. One, then, may argue whether this could be true also for moral conversions.

In what follows I will expand on Dees' philosophical analysis of moral conversion, which I believe helps us understand what it is that makes them *moral* conversions; then I will also expand on James and explore possible qualitative differences between moral and religious conversions.

⁸ Although he notes that "[l]ate in life [...]" in an article that was published posthumously (Kohlberg & Ryncarz, 1990), Kohlberg presented an unexpected twist on his theory by hypothesizing the existence of a further, metaphorically named, "Stage 7" [...]. Perhaps influenced by existentialist and humanist psychology (see e.g. Maslow, 1964), Kohlberg and his colleague now mused that there could be a unique peak experience stage of moral development, characterized by a "cosmic" or "transcendental" perspective, where moral truths are embraced through existential resolutions— intuitive jumps into the unknown— or sudden gestalt-switches, rather than dispassionate moral reasoning (see Gibbs, 2014 , pp. 91–93). Notwithstanding this evidence of a late Kohlbergian "conversion," it seems that the 1990 article has escaped the attention of most psychologists, and the most immediate arguments against epiphanic moral conversions in social scientific circles continue to be of Kohlbergian provenance" (Kristjánsson 2018: 20).

⁹ See for example Mark Jonas 2015a, 2015b and Asher 2017 on Platonic moral psychology and change.

1.2 Moral or Psychological?

“When I read Tolstoy's essays on non-violence and civil disobedience at age 13, I was so impressed by their moral strength that I began to question many of my political convictions. Reading those essays *awakened* my political consciousness, and I was led to *question* my unreflectively religious, firmly patriotic, Texas-style conservatism. In the long process of change that followed, my values – indeed, my deepest thoughts about value – were *transformed* until I had become a reflectively atheistic liberal. In a word, I had undergone a moral conversion” (Dees 1996: 531, my italics).

While introducing his discussion on moral conversion, Richard Dees provides us with the following elements: an occurrence (reading Tolstoy's essays) unpredictably triggered an awakening (of his numb political consciousness) and led to massive structural questioning (of his beliefs, as well as of his values), where what was questioned was then transformed into something radically different (a reflectively atheistic liberal outlook). Plus, such change was *unintelligible* for those around him. He then goes on to explain such change in causal terms:

“my views did not change *because* I had discovered a new truth about morality – though perhaps I had – but because I needed to create a psychological distance from my father. And while the discovery that the change was causal rather than moral might have undermined my faith in those values, it did not; I do not now regard those changes as any less justified because the best explanation for them is causal rather than moral” (ibid.).

Why did he change? The cause is exogenous: reading Tolstoy's essays. But not all Texas-style conservatives who read Tolstoy's essays end up undergoing a moral conversion. Moreover, it is not clear whether “moral truths” did any job, since Dees

himself explains it in purely psychological terms. In short, one may just say: right, you had father's issues and happened to read something that, were you to endorse it, would have allowed you to distance yourself from your father. So, you endorsed it. Of course it is radical, of course it is experienced as something dramatic, but did morality do any actual job here? To put it differently, do we need to call it a *moral* conversion? Is it enough that a conversion involves values for it to be called "moral", or do we need values to have a specific role in this conversion process to call it "moral" conversion? Is it enough that a conversion is "for the better" for it to be "moral"; that is, evaluated positively when it comes to morally relevant outcomes? Concerning this latter question, Kristjánsson's point 6 asserts that epiphanic moral conversion is a change for the better, thus suggesting that the "moral" of the conversion has a lot to do with how we externally judge this change. However, the same conversion can be judged in different ways; plus, it is not impossible to conceive of moral conversions for the worse.¹⁰ We need something that is moral not much in content (i.e., something either "good" or "bad"), but in its form, that is, that has to do with the relation between the agent and values; thus, with the first question.

Note that we can always find purely psychological explanations in conversion narratives, but this need not undermine the importance of the "moral" aspect of these radical changes. As stressed by Kristjánsson, whatever it is that triggers the conversion, it also discloses something to the agent: it may be different values, or a different way to look at the same values. It is whatever the agent sees, or experiences, or perceives, that has moral worth. This is why "epiphanic" moral conversions is a particularly apt label: epiphanies are "sudden disclosures of value" (Chappell 2022) which, by themselves, may not trigger anything, or nothing that lasts longer than the epiphany itself. But

¹⁰ Even if from the perspective of the convert, conversions will probably all be experienced as "being now on the right path," the convert can still be evaluated morally from an external perspective. See also note 16.

“epiphanic moral conversion” suggests that this sudden disclosure of value has triggered a radical change. Now, whether we can also find – as it is more than likely – perfectly reasonable psychological explanations as per why the agent was in the first place open – or resistant – to change, or why that particular trigger, does not undermine the fact that the trigger has disclosed something valuable that has changed the person in a radical way. This is why it is a “moral” conversion.

Now, Dees identifies different kinds of moral conversions (evolution, discovery and revelation), and when it comes to either of them, what really matters is why the agent felt compelled to change, or felt compelled to endorse a change that she – more or less – unexpectedly underwent. Thus, for each kind of moral conversion, he provides an argument for an internalist view of morality, where moral conversion is justified starting from the agent’s context.¹¹ Let us explore his conversion paradigms: *conversion by evolution, conversion by discovery and conversion by revelation*.

Conversion by evolution is exemplified by Dee’s own conversion as depicted above. He understands it as a gradual process in which he did not reject his values at once, but, little by little, he realized that a significant change had occurred, where “[e]ach change made sense in the context of the values I held at that particular moment” (1996: 534). In this case, justification is provided by the level of coherence between the convert’s core beliefs and values, and the new beliefs and values. If this is the case, then the conversion makes sense for her; and since such justification entirely depends on those values that are most important to her, in order to get to know whether this is actually the

¹¹ This is why Dees stresses that “[c]onversions do not begin with what the convert does, but with what happens to her. It requires a jolt: an event that forces her to re-examine her life and motivates her to change. [...] Her reasons for change only make sense to us once we understand the causal forces at work, so we can only judge those reasons once we understand the particular context of her conversion” (1996: 533).

case we need to look at her context. Thus, justification is plainly internal and contextual (535). However, even if the justification is purely internal, it still allows for criticism regarding the new values that are now important to her, as well as regarding the practices that follow suit.

Conversion by discovery is exemplified in the paper through the narration of second wave feminism's consciousness-raising groups. Once women belonging to traditional households began to talk to each other, they realized that theirs were shared experiences of something that was not right, and that "the very structure of traditional marriages made them vulnerable to the humiliation that they often suffered. What they once saw as a cozy and loving home life, they now saw as an oppressive and demeaning environment. Thus, the process of consciousness raising did not simply reveal hidden tensions in their lives; it changed their *experience* of their lives" (539). In this case, justification works in a way akin to that of a new scientific theory that can explain already available data according to a new perspective, that still validates the old data. Feminism, according to Dees, gave these women the power of legitimization of their past experiences, as well as a future direction. "The move is, then, rationally justified, in the manner Taylor describes: the new view has defeated the old, because it can explain both the successes and failures of the old, while the old cannot explain the successes of the new" (541). Justification is internal and contextual, since we still need to understand women's context in order to see how their old views got incorporated into the new ones.¹²

¹² It is important to remark the condescending tone of Dees' article, where he seems to consider women who attend consciousness-raising groups as children who learn new, eye-opening notions in a classroom. For a more accurate picture, see Kathie Sarachild's consciousness-raising manifesto (1973), also considered in Chapter II of this dissertation.

Conversion by revelation is quite the different story, where the peculiarity is that of an experienced collision with a new, overwhelming, moral outlook. When this kind of conversion takes place, even old standards of justification are transformed; it is a bit like a road to Damascus. Dees takes the example of Malcolm X endorsing the Nation of Islam's religious and moral outlook. "The Nation of Islam [...] presented the African-Americans who encountered it with a new vision of the world, one which did not simply clean up their beliefs and values or show them a truth about the world that forced them to take a new tack to achieve their values. Instead, it *revealed* to them a thoroughly new moral perspective, one which destroyed their old values and which challenged them to lead completely new lives" (544). The difference between conversion by discovery and conversion by revelation lies in the level of disruption between old and new moral outlooks: women in awareness-raising groups could look back and still understand why they were following patriarchal household rules, while Malcolm X, when looking back at his times in Harlem spent drug dealing was puzzled, he could not for the life of it understand why he was behaving as such. In some of these cases, the language of justification is not appropriate. Since the converted agent feels a discontinuity with her former self, she cannot justify her new beliefs in light of the old ones, as it is the case for the converted by discovery.¹³

I have stressed that it is the epiphanic element of disclosure of values that makes moral conversions peculiarly "moral". Does this imply that moral conversions only make sense if we assume that moral truth exist independently of us and, once discovered,

¹³ This is how Dees explains this element of incommensurability: "[t]he choice between the new values and the old is made on the basis of values that are themselves a part of the transformation, so we have no impartial standpoint from which to assess the conversion and we have no way to assess it rationally. In these cases, then, the contextualist view I have used in the first two kinds of conversions gives us no means by which to judge the change itself. We are left with nothing more to say" (1996: 546).

impose such a force on us we cannot but abide? Realism¹⁴ seems to find some support in the phenomenology of this kind of conversions: the agent experiences these radical shifts as a revelation of “a previously-hidden moral truth, which she feels compelled to accept” (1996: 548), so it seems that in order to justify such shift would imply for the agent to believe that there must be some sort of absolute perspective on moral issues, and that she can finally see it through her new moral outlook. Any counter-conversion would be a betrayal of that new truth. However, argued Dees, “the absolutists get the causal story backwards: they assume that the convert does not want to go back to her old view because she believes that her new view is validated from some absolute point of view, but in fact, she comes to believe that it is absolutely validated because she already knows she cannot return” (549).

This is tricky. Surely there must be a way to reverse the conversion, if the agent decides that her old values were, after all, better, or more important to her than the new ones. Dees does not deny this and ends the paper with the following suggestion: we should give up looking for absolutist justifications. “The convert must simply give up the idea that her new view must be validated absolutely. [...] Nothing more is needed” (ibid.). Again, all we need is to understand her context and, crucially, her motivations to either pursue or try her best to reverse the conversion. We do not need an externalist view of morality to make sense of any of the above-mentioned kinds of conversion, because even if we simply focus on the agent’s inner changes as they make sense for her, we can still find ways to justify or question her conversion. Justification must be connected to the agent’s motivation and, to do so, a central role needs to be given to people and their contexts.

I have isolated the following elements of moral conversion: it makes sense for the agent given her own context, it comes as an abrupt turning point, it is experienced as

¹⁴ Which Dees calls “absolutism” (1996: 547).

exogenous and it triggers a radical reconfiguration of the self, or of one's attitudes, or of one's action-guiding values. Moreover, what it seems to be the uniquely "moral" element is the fact that when a moral conversion happens, the agent is somehow forced to engage with the morally-laden world in a way that is radically different from before. This can take the form of a gradual adjustment to values that have been disclosed in virtue of the conversion. We are assuming one particular way to understand morality, which is the one that does not endorse the fact/value distinction and that, thus, sees the world as naturally value-laden. However, moral conversion can be uniquely moral even according to different understandings of what morality is. Alfredo MacLaughlin has provided a thorough analysis of different ways of making sense of moral conversion, depending on one's account of morality, to which I now turn.

1.3 The 'Moral' of Moral Conversion

In his doctoral dissertation "Narratives of Hope: A philosophical study of moral conversion" (Chicago Loyola University), MacLaughlin introduces moral conversion as a change in "deep-seated convictions, attitudes and patterns of action regarding moral matters" (MacLaughlin 2008: 1) that happens in "rather unexpected and surprising ways" (ibid). Such process has both an internal and an external aspect: it is a drama that is "played in an internal field of convictions, beliefs, attitudes and emotions, directly accessible only to the subject who is undergoing this process, and, if it becomes known at all to external observers, it is only through acts of communication of the convert" (2008: 10). The external element has to do with the fact that conversion affects the agent's attitudes and behaviors. It is, thus, to some extent observable. He goes on to specify the phenomenon as an "existential change" (2008: 77), since (i) it affects "that person's life fundamentally" (2008: 78), (ii) it has impactful consequences, (iii) it is

personal (as opposed to structural, institutional ...), (iv) it is not planned, but rather unexpected, and (v) has an emphasis on freedom (“existentialism”). Finally, he distinguishes three general “classes” of conversion: “conversion regarding content, conversion regarding attitude or degree of commitment, and conversion regarding behavioral coherence” (2008: 81). These classes intertwine with two ways of understanding morality:¹⁵ (RW) morality as the consideration of what is right or wrong, and (ML) morality in its classical notion, where it has to do with the quest for happiness, eudaimonia, or “the meaning of life”.

In the context of RW, moral conversion becomes “a change in the [moral agent] with regard to her/his existential involvement in the task of acting rightly or wrongly, of doing what ought to be done” (2008: 83). In particular, we can distinguish the following three declinations of moral conversion, where morality is understood as the consideration of what is right or wrong:

(RW1) Moral Conversion regarding the *content* of right or wrong. In this case, in order to recognize whether there has been a moral conversion, one needs to see whether the agent’s judgments have shifted in content, and whether such shift can be attributed to “deeper changes affecting a person’s overall structure of moral judgment, rather than changes in particular matters when the bases of such judgments have remained unchanged” (2008: 85). According to this picture, a change in decisional criteria would

¹⁵ In *Moral Teleology* (2023) Sauer also distinguishes between two sense of morality when it comes to changes for the better (moral progress): the narrow sense, whereby “morality is *the thing that undergoes the improvement*” (28), and the wide sense, whereby “[t]here is some other form of social change or development which is *welcome* from the moral point of view, without morality itself necessarily being the thing that improves” (ibid). One could also understand MacLaughlin’s distinction as one that sees, on the one hand, improvements in morality understood as “good changes,” independently from one’s favorite moral theory, and on the other improvements in morality understood as “whatever it is in one’s life that is morally significant.”

count as a moral conversion. A person who is skeptical about climate change may begin to judge that it is indeed a serious issue, but that would merely be a change in judgment, or opinion. If, however, she begins to consider her own actions as morally relevant, not only she may change her opinion, but also her actions; and if she includes considerations about her own ethical impact in her decisions, this would amount to a moral conversion, as opposed to a mere change in judgment (2008: 84-85). It is, thus, not only a change in content of right/wrong considerations, but one that happens due to a change in the agent's decisional apparatus. In short, it is a sort of *methodological* conversion regarding what grounds moral decisions.

(RW2) Moral Conversion regarding the *degree of commitment* with respect to right/wrong, which can be described as “a shift from a general lack of concern for the morality of one's actions to genuine concern for this” (2008: 90). In this case, there may be also a change in content, but the significance of the change is to be found in the agent's *attitudes* towards morality. A case identified by the author is that of Gandhi being forced out of the train: when he was still a representative of his Indian law firm, he was travelling first-class on a train in South Africa, when an official told him to move to third-class due to the color of his skin. Upon refusal, he was forced out of the train. While sitting in the waiting room, he realized he could have continued his journey, or he could have dedicated himself “to root out the disease and suffer hardships in the process” (Gandhi, *An Autobiography*, 1927: 103-105). This, according to MacLaughlin, can be described as a change in attitudes and commitments towards morality.¹⁶

(RW3) Moral Conversion regarding *behavioral coherence* in right/wrong., where the agent both knows what she should do and acts on such knowledge. This case may be

¹⁶ There is also the possibility of “counter-conversion”, where the attitudes and commitments towards morality change for the worse. This is exemplified, for example, by the narratives of Patrick K. and Robert Cooley (100-101). However, they are more akin to moral “downfalls” than to proper counter-conversions.

assimilated to standard virtue ethicist moral development, where the agent develops into a mature, virtuous agent through the consolidation of good habits.¹⁷ However, according to MacLaughlin, it may still be possible to identify a shift in the agent that is not due to standard moral development, but rather to a conversion, where “the person realized that his/her commitment to his/her own moral standards is weak, because of an apparent lack of coherence between the actions and the contents of the moral standards[, then] the person judges this low degree of commitment to be unacceptable, and commits in earnest to certain moral standards, and accordingly begins acting more coherently” (2008: 104). In order to distinguish this type of conversion from RW2, MacLaughlin adds that since a greater level of commitment can still coexist with behavioral incoherence, a conversion has happened when coherence is spontaneous and not a struggle.¹⁸

If, on the other hand, we consider ML, then moral conversion turns into a slightly more nuanced process (although complementary of RW, 2008: 120), which has to do with questions about life, such as “the issue of what kind of life is worth living, or what kind of life would fulfill the promise implicit in my particular talents, or the demands incumbent on someone with my endowment, or of what constitutes a rich, meaningful life ...” (Taylor 1989: 14). Combining the three dynamics of moral conversion with ML, we get the following picture.

(ML1) Moral Conversion regarding the content of happiness/eudaimonia/the meaning of life, where the agent shifts from looking for meaning in one category of goals, to looking for it in a different one (MacLaughlin 2008: 124). Categories of goals are those categories where the agent looks for the answers to questions about eudaimonia (e.g.

¹⁷ More on this in Chapter III, section 3.

¹⁸ This, I believe, can still be easily reduced to standard moral development in virtue-ethicist terms. I will expand on this point in Chapter 3.

possessions, entertaining, hobbies, creative expression, knowledge acquisition, etc.). Just like in RW1 a mere change in opinions does not amount to a moral conversion, here a shift of preferences within the same category would not amount as a moral conversion. That is, if I look for life fulfillment in material possessions, changing the kind of material possession does not amount to the kind of change that is required for a conversion, since it is not “structurally significant” (2008: 127). On the other hand, shifting from one category of goals to another may be considered an instance of a moral conversion. It can also be exemplified as a shift in the *questions* about happiness/eudaimonia/the meaning of life, instead of in the answers, or as a dismissal of questions of eudemonia, as the author understands the case of Frodo in *The Lord of the Rings*, where questions of duty became much more salient than his own happiness (2008: 135).

(ML2) Moral Conversion regarding a change in *attitude* towards happiness/eudaimonia/meaning of life. In Socratic terms, we can understand it as a shift from an unexamined to an examined life. There seems to be a variety of attitudes towards ML, such as “optimistic acceptance, joyful awakening, grateful relief, blind obsession, even an anguished struggle to fend off despair, to escape nihilism ...” (143). Only some of those attitudes can be considered “good” in relation to ML. When a person’s attitude has changed “for the better” and has become a habit, we may consider it a form of moral conversion.¹⁹

(ML3) Moral Conversion as a change towards more coherence in one’s behavioural patterns, where the agent attempts to conform her patterns to the requirements of eudaimonia. This type of conversion is experienced through “an infusion of new strength, new enthusiasm, new operative hope, on the wings of which the person is able

¹⁹ There is of course the possibility of counter-conversions, where the idea is that of actively seeking an unexamined/uncritical/merely factual life is paradoxical (e.g. Sartre’s *Nausea*).

to overcome effectively the obstacles and contrary habits that weighed him/her down, and engage in ... patterns of behavior that are consistent with what is regarded as the proper direction towards happiness/eudaimonia/meaning” (157). The crucial aspect here is the focus on eudaimonia within conversion narratives. MacLaughlin introduces Bob, a doctor and an ex-alcoholic who, when narrating his case, mentions drinking as an obstacle to his flourishing rather than as a matter of right/wrong, or religion. His strong desire and hope to get rid of this obstacle – along with a number of other, intermediate steps²⁰ – is what allowed him to be coherent in his behavioral pattern towards what he deemed a flourishing life; namely, one in which he was not an alcoholic. This shift is, thus, according to the author, a moral conversion.

We can schematically summarize MacLaughlin analysis of the combination between dynamics of moral conversion and two different ways of understanding morality: RW, where morality has to do with what is “right” or “wrong”, where facts and values are sharply distinct, and ML, where morality has to do with the meaning of life, and there can be room for both a value-laden world and a world of natural facts that is distinct from that of values.²¹

²⁰ Talking to friends, but also being assisted by another ex-alcoholic (MacLaughlin 2008: 160).

²¹ Famously, Moore has advocated for the necessity of a fact/value distinction in virtue of a non-naturalist account of the term “good” (1903).

	RW: Morality as right/wrong	ML: Morality as the search for meaning in life
Content	RW1: changes in judgments of what is right/wrong and in decisional criteria. “Methodological” conversion.	ML1: changes from looking for meaning in one category of goals (e.g. possessions, hobbies) to another (e.g. pursuit of knowledge, social justice).
Commitment/attitude	RW2: from a lack of concern regarding what is right/wrong to a concern of what is right/wrong.	ML2: from an unexamined to an examined life.

Behavioral coherence	RW3: change from forced or no coherence between commitments and actions to spontaneous coherence between commitments and actions.	ML3: change in conformity between behavioral patterns and the requirements of eudaimonia.
-------------------------	---	---

In both the case of RW and ML, the uniquely moral element involves changes in judgments, considerations, level of commitment, attitudes towards what we deem right or valuable and coherence between what we preach and what we do. These are descriptions of how these changes can go, of the possible paths of radical change that the convert can take. These descriptions are compatible with what we observed about epiphanic moral conversion: that is, that what it is that makes them uniquely moral is that all these possible paths need to be *disclosed* somehow. Thus, if we combine Kristjánsson, Dees and MacLaughlin we can see that moral conversions can be understood as

radical moral changes triggered by (more or less) sudden disclosure of value that leads to a variety of possible patterns of change.

We have successfully isolated the “moral” element. What is left to analyze is whether there is any substantive difference between a moral and a religious conversion; it is to this question that I now turn. I will first consider religious conversion, then argue that we have good reasons to understand it as distinct from moral conversion.

1.4 Moral and Religious Conversions

In her book *Turning* (1980) Emilie Griffin proposes a four-stage model of the phenomenology of religious conversion, grounding her analysis on some famous conversions, such as those of Augustine and C. S. Lewis, as well as on her own experience. The stages that the converter undergoes are: desire, dialectic, struggle and surrender.

Desire. This stage is characterized by a general feeling of dissatisfaction, which can be experienced by people with different degrees of life accomplishments. One can be an incredibly rich CEO, with the perfect family and a vibrant social life, yet feel the desire, the longing for something else. Griffin reads this stage as the natural longing for God that is present in each person, but she explains that, while one is undergoing such stage, one does not experience it as a *clear* desire for God, but rather as a mysterious longing for something or someone that cannot be called yet. It is a stage characterized by the difficulty (and, possibly, the fear) of giving a name to such “something/someone”, along with the intellectual drive to solve the mystery (Griffin 1980: 48-49).

Dialectic. The second stage is also called the “argumentative” phase, where the agent pretty much spends her whole time arguing against the possibility of calling this

“something/someone” with the name ‘God.’²² It may also take the form of an argument about pros and cons of endorsing God’s will, or as the search for possible alternatives. It is a phase of initial resistance, but one that can still be somewhat pacific; it may be exemplified by the image of two educated intellectuals trying to find middle ground on incompatible viewpoints. It may also have an element of confabulation,²³ since the agent tries to reason about spiritual impulses that belong to a completely different logic than the one she tries to employ. The only way to solve this situation is with the “leap of faith,” a jump beyond logic and reason.

Struggle. As explained by MacLaughlin, in this phase “one foot is already on faith ... but the other foot is ready to step on the breaks” (2008: 65). It is a stage of more or less violent war against the new possibility, with which the converter was merely arguing just one stage before. It is interesting to note the possible reasons behind this resistance. The agent may begin to see that she is already half-way through surrendering, but she may recognize the difficulty all this would imply for her life (e.g. other people may not understand, it may clash with the idea others have of her, a general “fear of change,” 1980: 123). Griffin also considers possible feelings of unworthiness, where the agent feels she is not ready, or that she is too lacking, to endorse God (1980: 110).

Surrender. Finally, the agent fully assents to whatever she is converting to (as in *Surprised by Joy*, Lewis 1955). This moment of “letting go” (MacLaughlin 2008: 68) is accompanied by the experience of total exhaustion, which does not take place “until the convert somehow gets beyond wanting, beyond his own efforts and demands, into a state where he wants nothing or cannot say with any clarity or confidence just what he wants. At this moment when he wants nothing, he is able to will something – but

²² Note that she specifically considers Christianity.

²³ Whereby convert may remember or believe things that are actually not true, but functional to maintaining an integral picture of one’s self (Schnider 2008).

dispassionately: in Milton's phrase, 'all passion spent'" (Griffin 1980: 143).²⁴ All that is left is abandon ones' self to God's will.

Now, Kristjánsson considers Griffin's analysis as well, and notices that

"the more accurate a picture we take this model to paint of religious conversions, the more those seem to differ from moral conversions, or at least the sort of moral conversions under scrutiny in this chapter [epiphanic moral conversions]. For, as MacLaughlin (2008 , chap. 4) correctly points out, the latter are typically less reflective than the conversions identified by Griffin; they do not universally present any preliminary desire phase; and they often include no experience of a struggle and eventual surrender." (Kristjánsson 2018: 28).

In other words, Kristjánsson – following MacLaughlin – notes that conversions that are uniquely moral do not share the same phenomenology with those that are uniquely religious. At least, not those he is considering; that is, epiphanic moral conversions. One does not need to be dissatisfied with one's moral status quo in order to be struck by a conversion; that is, many narratives of moral conversions do not involve the agent actively desiring, or longing for other possible values to be guided by. Moreover, narratives of moral conversion do not often involve moments in which the agent fights against the possibility of being converted to new values and, finally, surrenders to them. In short, it is not so easy to replace "God" with "new values" or "a new moral theory". As MacLaughlin puts it, "religious conversion has the form of a search or quest, and there is the drama produced by the existential involvement of the whole person in this

²⁴ William James on the "psychology of self-surrender" (The Varieties of Religious Experience, p. 173).

quest, while moral conversion very often has rather the form of an awakening, or a slap in the face” (2008: 74).²⁵

On the other hand, in both kinds of conversion there seems to be some sort of change in the person (e.g. relations of the person to the world, to others, to truth, to the self, how the person is affected by such change, etc.),²⁶ which is what MacLaughlin means when he claims that it is an “existential” change, that has to do with persons and not to, say, institutions or systems of laws. Moreover, the threefold notion of conversion concerning content, attitude and coherence can be found also in religious conversion (2008: 165 – 170). Thus, he concludes, even though the phenomenology of religious conversion is not entirely shared with that of moral conversion, there are still some significant overlaps. As we mentioned at the beginning, a religious conversion is likely to be also a moral conversion (Leuba 1896) while a moral conversion need not imply also a religious conversion. We can imagine a person who is already morally straight, acting well just for the sake of duty, or due to a general respect for morality, who converts into Christianity and, thus, finds another – this time spiritual – purpose to believe and do what she already believes and does. According to MacLaughlin this would be a religious conversion in content; that is, in the structure of one’s grounds for action. However, it is not a moral conversion, since a concern for morality was already present

²⁵ MacLaughlin does not seem to consider the fact that one may undergo proper moral crisis, and not just experience a slap in the face. Again, that of Jean Valjean in *Les Misérables* is a beautiful depiction of how a moral awakening can be experienced as painful and worth fighting against, which suggests that moral crisis do share the same phenomenology of those religious crisis that precede the conversion.

²⁶ There are also intellectual conversions, understood as “the redefining of one’s understanding of notions such as certainty, facts, knowledge, etc., in such a way that the person’s cognitive relation to the world is fundamentally altered” (MacLaughlin 2008: 77).

in the structure. What is it that – if at all – distinguishes the two kinds of conversion, given that the phenomenology seems to be different, but need not be so?

It has been recently acknowledged that the phenomenon of religious conversion lacks philosophical consideration (Spinney 2023). It was considered from a psychological perspective, famously by William James (1902), and it is currently studied through the lenses of neuroscience, history, anthropology and literature (see *Oxford Handbook of Religious Conversion*, 2014). This lack of philosophical conceptualization makes it so that the label “religious conversion” is, at best, vague (Spinney 2023: 140). There is no consensus over a definition of religious conversion, nor on how to spot the truly convert as opposed to one who merely endorses or adopts a new religion without undergoing a proper conversion. Still, we can attempt to find at least a working definition of religious conversion that can help us distinguish it from that of moral conversion.

William James (1902/1982) defines conversion as “the process, gradual or sudden, by which a self hitherto divided, and consciously wrong inferior and unhappy, becomes unified and consciously right superior and happy, in consequence of its firmer hold upon religious realities” (150). This definition has been challenged by Griffin, who claims that religious conversion does not necessarily involve a divided self that, after the conversion, is reunited, nor one that finally reaches happiness (1980). As a matter of fact, the convert may feel relieved due to her surrender, but it does not imply that she is going to be happy about her radical change. Quite the contrary, the convert’s life is likely to be much harder than it used to be. Griffin is not satisfied with a purely psychological account of religious conversion, and I believe it to be due to the following reason: psychology can describe the “conversion” experience, but not the *religious* one. Indeed, James provides a thorough explanation of the conversion process, where the gravity centre of the agent, her “hot place”, radically changes. He defines the “hot place” of an agent as “*the habitual centre of his personal energy*” (155); that is, the place

where we can find those things for which he “burns”. In other words, we can say that the hot place is where we can find the agent’s action-guiding motivations, her drive to act. James goes on to say that “[i]t makes a great difference to a man whether one set of his ideas, or another, be the centre of his energy; and it makes a great difference, as regards any set of ideas which he may possess, whether they become central or remain peripheral in him. To say that a man is “converted” means, in these terms, that religious ideas, previously peripheral in his consciousness, now take a central place, and that religious aims form the habitual centre of his energy” (ibid.). This is hard to articulate for both the external and the internal observer. “We have a thought, or we perform an act, repeatedly, but on a certain day the real meaning of the thought peals through us for the first time, or the act has suddenly turned into a moral impossibility” (155-156). All we know, all we can say is that some things are now “cold”, dead, and others have become “hot.”²⁷ Mental rearrangements are like house re-building: when the gravity centre changes, the building has to be re-structured around it. Such changes in gravity centers may be due to sudden and violent emotional reactions (e.g. “love, jealousy, guilt, fear, remorse, or anger”: 157).

This is a depiction of the process of any conversion – but what is it in religious conversion that makes it truly *religious*? From an external perspective, one can just observe the behavioral changes in the convert where, say, she used to go to a Catholic Church and now she goes to a Mosque, and legitimately say “she must have converted.” This would explain her change in behavior, but does “religious” merely mean “religious system”, or “set of rituals”? If that was the case, then it would be like saying that a person has morally converted if she used to vote Republican and now she votes Liberal. It is indeed a change, but we can say no more. Was it gradual? Was it a mere change of

²⁷ See also Caprioglio Panizza and Vaccarezza (2024) and Liberti (2024) for a similar point.

opinion, or perhaps a brainwashing? All we have seen so far regarding religious (and moral) conversion involves the inner world much more than the external one. Thus, it is there that we should look for the peculiarly religious element. James recognizes that the common element of religious conversion's narrations is the experience of lack of control over what is happening, and the consequent realization that true religious conversion is an expression of a Deity's will (1982: 190). This is why one has to surrender to it. This is what, according to Griffin, psychology cannot explain, although James tries his best (e.g. when introducing the concept of "automatisms" as ordinary expressions of the subliminal mind: 184). I believe that James offers an incredibly intelligent reading of what psychology cannot ultimately explain: he does not deny that Deity might be truly at work in religious conversion – thus making it not a purely psychological occurrence – but if Deity is at work, then it operates through what is "Subconscious" (190). This seems to imply that, since we are not all the same when it comes to the demarcation between what is conscious and what is not inside of us,²⁸ then there will be agents more prone to religious conversion, and agents less prone to it. When there is no mysteriously subconscious activity involved, then it is just an ordinary conversion in the terms understood above: a reconfiguration of one's gravity centre which, in James' terms, is a perfectly normal psychological occurrence in people's lives. In short, James leaves room for the possibility of both the existence and the direct action of Deity on human life, at the same time saving the psychological underpinnings of conversion experiences.

²⁸ As James quotes, Professor Coe's study on religious experience (1900) supports the following conclusion: "if you should expose to a converting influence a subject in whom three factors unite: first, pronounced emotional sensibility; second, tendency to automatisms; and third, suggestibility of the passive type; you might then safely predict the result: there would be a sudden conversion, a transformation of the striking kind" (1982: 189).

If all this is true, then we are left with the following possible distinction between religious and moral conversion: in religious conversion there is room for the inexplicable action of a Deity, while in moral conversion there is no room for the inexplicable action of a Deity. That is, in both cases there can be – and, as mentioned before, there likely are – perfectly intelligible psychological reasons that explain why the agent was triggered in one direction rather than another (as stressed by Dees 1996). However, in both cases there is room for something that is not so psychologically intelligible, what James calls the “subconscious”. There is a mystery that can be experienced in different ways: when it is experienced as the direct action of a Deity, then we call it a “religious” conversion; when it is experienced as the disclosure of values that significantly shake and re-shape the agent’s moral outlook, then it is a “moral” conversion. As we mentioned, one does not exclude the other; moreover, one can temporally precede the other. One can have one’s moral outlook shook and re-shaped due to the experience of a direct action from a Deity; thus, experience a moral conversion *in virtue of* a religious conversion. One can experience a sudden disclosure of value and, upon undergoing a massive revision of her moral outlook, experience also a religious awakening that she attributes to God’s direct call; thus, experience a religious conversion *in virtue of* a moral conversion. However, in a conversion that is uniquely moral, one need not experience any direct intervention from a Deity of some sort. The mystery of the uniquely moral conversion is just that no matter the amount of perfectly intelligible reasons – moral, intellectual, psychological – that the agent may have to be both open to the epiphany and triggered towards one particular path, it is still impossible to make full sense as per why, of all the things that have happened to her, it was *that* one which triggered the epiphany; and of all the paths that the trigger could disclose, she went in *that* specific direction. James’ analysis helps us in distinguishing moral and religious conversion when considering the “subconscious” part: in religious

conversion, the subconscious is the locus of – the possibility of – Deity’s direct action; in moral conversion, the subconscious is the locus of the epiphany. This does not undermine the uniquely moral element, which is the disclosure of value and its subsequent pervasiveness in one’s attitude towards the world.

In conclusion, we can define a moral conversions as

radical moral changes triggered by (more or less) sudden disclosure of value that lead to a variety of possible patterns of change

where the uniquely moral element includes both the attitude towards and the content of morality, be it understood as guiding us to what is right, or to what is good, and can be, to some extent, distinguished from religious conversion in it not needing a Deity in order to be intelligible. Finally, by analyzing existing accounts of moral conversion, we can notice that they all establish, more or less explicitly, the following common elements: a moral conversion is experienced as abrupt and unplanned, it triggers a radical reconfiguration of one’s moral outlook and, for this reason, it comes about as disruptive of what came before. It is now time to turn – forgive the pun – to Moral Transformations and see what it is, if anything, that they may have in common with the phenomenon we have just analyzed.

2. Ladders of Moral Transformation

‘Transformation’ is composed of the Latin prefix ‘trans’ which means ‘across, beyond’ and ‘formare’, to form; thus, it involves a process of going beyond a certain form, to undergo a change in form. This suggests that a “moral” transformation may involve a morally significant change in form. In what follows I will expand on two ways to read the phenomenon of moral transformations: the “ascension” reading, and the “activity” reading. The “ascension” reading sees its origins in Plato’s ascension passages in *The Symposium*, and had been taken seriously in more recent times by Wittgenstein and Murdoch, who understand moral transformations climbing the steps that separate us from contemplating the Ideal – be it Beauty, the Good, or perfection. The “activity” reading has been recently advanced by scholar Agnes Callard who, in a diametrically different fashion, sees the process of moral transformation as being ultimately practical, rather than contemplative. One could say that this whole section is nothing but another philosophical depiction of Raphael’s *School of Athens*.

2.1 Transformation as ascension

What if that ‘Deity’ we mentioned before and ‘value’ *were* the same thing? I mean, not exactly the same thing, but what if they belonged to the same category; namely, what Wittgenstein called “the higher”, that of which nothing can be said? If that was indeed the case, then moral ascension would also be a mystical ascension. This idea is present in Plato, and has been later revived by Schopenhauer, Wittgenstein and, in a different fashion, also by Murdoch. These philosophers had the intuition that contemplation may

trigger moral ascension; that is, that truly seeing something means seeing it as it is, and that seeing something as it is, is a transformative phenomenon.

Ascension is exemplified by the analogy of the ladder: in the *Symposium*, lovers climb up to the highest level of love they can reach (209e5–212a7); reading the *Tractatus*, we start from meaningful propositions about states of affairs in the world, and we climb up to higher states of affairs, of which we cannot speak (TLP 6.54) – and then, we kick the ladder away. When it comes to Murdoch, the metaphor relates to the metaphysics rather than to the phenomenology, but we can still see an ascension path towards the endless task of truly accessing reality (IP 317).

2.1.1. The *Symposium*

In the *Symposium* we learn that Eros, far from being a perfect god, is a creature of desperation, conceived from rape, who constantly seeks to relieve his neediness by chasing after the beautiful. Being in a state of eros, thus, implies being aware of one's own imperfection and experiencing neediness for perfection. In Plato's lovers ladder, those lovers who seek beauty for the sake of becoming immortal are the lowest ones (psychic lovers, Obdralzek 2010: 10), since such kind of love is completely self-absorbed, and the other is used entirely in order to secure one's own illusion of immortality; they lack true virtue, since they lack true knowledge of the Beautiful²⁹ (kalon, Obdralzek 2010: 13). In the "ascent passage" (209e5–212a7), on the other hand, we see a different eros, one that is transformed through the love of beauty; that is, "we

²⁹ Although Diotima says the exact contrary. However, we have good reasons to doubt that Diotima is the voice of Plato (Obdralzek 2010), at least if we take her speech in itself and not in relation to the others, especially to that of Alcibiades (Destrée 2017).

see what it is for ordinary *erōs* to be pursued correctly and transformed into something sublime” (2010: 17). In the lower stages *eros* seeks only beautiful bodies, then “he loves the man even more if he has a beautiful soul (209b4–7); and he gives birth to *logoi* about customs and virtue and to laws (209b8–c1, 209d6–7)” (2010: 17). Lower *erotika* first seek bodies, then soul, then customs and laws. Higher *erotika* follow the same path, only they do it *correctly* (2010: 8), so that they can finally see Beauty. Higher *erotika* do not make the mistake of thinking that Beauty is the means to some other end; rather, they seek Beauty as a *telos*:

“What he [Socrates] hopefully comes to recognize is that *erōs* does not have a use, and that [[beauty]] is not a tool to some other end; the contemplation of Beauty has intrinsic worth, and should be one’s life’s pursuit. To see this, Socrates has to radically rethink what beauty is and what our proper relation to it is. The [[beauty]] in question is not mortal, but divine, and it is not something to be possessed, but rather admired” (2010: 19).

Another error of the lower *erotika* is that of seeking immortality, because in seeking immortality they see love as possession.³⁰ However, the Forms are not the kind of things that can be possessed;³¹ in the ascent, possession is transformed into contemplation and admiration. Thus, the highest form that a human life can take is that of contemplation: “To contemplate the Forms does not make us perfect—only Forms are—but by engaging in an activity which ties us to the best objects, we become as perfect as humans can be” (2010: 21; Destrée 2017).

Transformed *eros* is not only *eros* that contemplates Beauty, but also *eros* that, in order to contemplate Beauty, is driven away from himself. The higher lover is one who is not

³⁰ See the passage 206e8–207a2.

³¹ See Destrée (2017) for an interpretation that sees happiness of the wise as the “eternal” possession of true virtues, gained by their ascent to the Forms (“*eudaimonist*” reading of Plato).

self-centred, but whose gaze is focused on abstract Forms; it is more objective, since a self-centred gaze is one that is far away from the true Forms and, consequently, distorts reality. Transformation through (and towards) Beauty is, thus, a process of radical decentralization. Let us see how this unfolds:

“The initiate is focused on observing and admiring [[beauty]], not on possessing it. The result is a kind of self-forgetting: he is no longer aware of his imperfection, because he is caught up in the perfection of the Form. This involves *a radical shift in the initiate’s priorities*, one which is captured by Plato’s use of the language of mystical initiation. By contrast, when one is focused on possession, one sees the beloved in terms of his relation to oneself, and remains self-absorbed. Similarly, when one treats the beloved as a means to immortality, one’s attention is turned towards one’s own defective, mortal nature” (2010: 23, my italics).

Even though higher lovers do not become perfect themselves, in focusing their attention to perfect entities they forget about their own imperfections and, thus, cease to seek for perfection in the non-ideal world. In the higher lovers, eros is transformed into a “constructive and selfless desire for the contemplation of Beauty” (2010: 24).

This is not the only way to interpret the Symposium; as a matter of fact, Kraut (2017) points out that there is also what he calls the “eudaimonist” reading, according to which the final goal of ascent is that of one’s own happiness (Destrée 2017), which translates as “immortality” insofar as contemplating the Forms provide one with eternal virtue which, in turn, provides all the goods that a happy life needs (as Socrates states in 30b). According to this reading, that of ascent is not a process of de-centralization, and, thus, may not even count as a proper transformation, since the agent seems to be the same through and through. In other words, through ascent one gets closer to the Forms, but does not get transformed by them; the “selfless desire for contemplation” becomes the active desire for wisdom, achieved through proximity to the Forms. However, according

to Kraut, we have good reasons to believe that the “eudaimonist” reading is not the one that Plato intended, since it clashes with his guiding ethical thought, which is that “one should live one’s life in response to something superior to one’s self” (2017: 236), where the agent, in her ascent, ceases to be important.³²

This is not the place to dispute the correct interpretation of the Symposium, and of Plato’s philosophy at large. What interests this research is the fact that philosophers who took the phenomenon of moral transformations seriously were inspired by two Platonic themes: (i) ascension as a process of moral transformation and (ii) contemplation, which is what allows moral transformation. We can find both (i) and (ii) in Wittgenstein’s *Tractatus*,³³ and a version of both in Murdoch.³⁴

2.1.2 Ascension and Contemplation in Wittgenstein’s TLP

In TLP we find a series of striking ethical claims:

6.4 All propositions are of equal value.

6.41 The sense of the world must lie outside the world. In the world everything is as it is, and everything happens as it does happen: in it no value exists—and if it did exist, it would have no value. If there is any value that does have value, it must lie outside the whole sphere of what happens and is the case. For all that happens and is the case is

³² We are unimportant non only in relation to the Forms, but to other external realities such as the polis. This is emphasized, for example, in Socrates’ speech in the *Crito*, where he concludes that he cannot escape from jail since his life is less important than the will of the laws of Athens (50b, 50e).

³³ On the Plato-Wittgenstein connection see von Wright, p. 19, from Malcolm’s L. Wittgenstein, A memoir.

³⁴ See Ollson (2018) and Caprioglio Panizza (2022).

accidental. What makes it non-accidental cannot lie within the world, since if it did it would itself be accidental. It must lie outside the world.

6.42 So too it is impossible for there to be propositions of ethics. Propositions can express nothing that is higher.

6.421 It is clear that ethics cannot be put into words. Ethics is transcendental.

(Ethics and aesthetics are one and the same.)

Morris (2008) draws the argument contained in 6.41 as follows:

(V1) Value is non-accidental;

(V2) Everything in the world is accidental; so

(V3) Value is not in the world.

He then notes that “... it seems clear that Wittgenstein’s conception of ethics involves some kind of removal of the ethical subject from the world.”³⁵ And this provides what seems to be an independent source of (V3), the claim that value is not in the world” (2008: 325). This provides the first element for the ascension: it seems that if we stay in the world, everything we can say about value will be meaningless, since we could only utter meaningful propositions about facts of the world. But if we move *outside* of the world we may have the possibility of saying something meaningful about value.³⁶ Moreover, it seems that in order to be able to say something meaningful about value, we

³⁵ A fact-value distinction which, on the other hand, is not found in Murdoch.

³⁶ V2 is defended by Wittgenstein in 6.373 and 6.374, where he endorses Hume’s rejection of natural necessity (Morris 2008: 321), while V1 is a rejection of consequentialism (TLP 6.422; Morris 2008: 322).

need first and foremost – and, possibly, only – to look at it in the *right* way, which is the second platonic element, that of contemplation.³⁷ This is why Wittgenstein claims that:

6.421 Ethics is transcendental. (Ethics and aesthetics are one.)

Four Notebooks entries and a further quote from TLP can help us understand what Wittgenstein has in mind when he equates ethics and aesthetics:

The work of art is the object seen sub specie aeternitatis; and the good life is the world seen sub specie aeternitatis. This is the connexion between art and ethics. (NB: 83)

(The thought forces itself upon one): The thing seen sub specie aeternitatis is the thing seen together with the whole logical space. (NB: 83)

Aesthetically, the miracle is that the world exists. That there is what there is. (NB: 86)

Is it the essence of the artistic way of looking at things, that it looks at things with a happy eye? (NB: 86)

6.45 The contemplation³⁸ of the world sub specie aeterni is its contemplation as a limited whole. The feeling of the world as a limited whole is the mystical feeling.

³⁷ 6.423 Of the will as the subject of the ethical we cannot speak. And the will as a phenomenon is only of interest to psychology.

³⁸ Here Morris notes that “ ‘view’ would be better than ‘contemplation’ as a translation of “ ‘Anschauung’ [since] it is an attitude, rather than an activity.” (2008: 325 – 326).

Thus:

- (a) the work of art and the world are equal in their being the thing they are when seen in a certain way; that is, when seen “sub specie aeternitatis”;
- (b) seeing something sub specie aeternitatis implies having a certain outlook: one that can see what it is that exists, how its part can combine with each other, and where there are no more possible combinations – the world with its possibilities and its limits;
- (c) the outlook of seeing something sub specie aeternitatis is one that is able to see that the very existence of the world is a miracle;
- (d) this particular outlook is “the happy eye”;
- (e) this “happy eye” gaze allows one to see and feel the possibilities within the limits.

Seeing the world sub specie aeternitatis is the element that is common to both ethics and aesthetics. In Morris’ words:

“There is a single ethical imperative, which is to view the world in a certain way. This is the artist’s way of viewing the world, a way of looking at things ‘with a happy eye’, an outlook which finds it a miracle that the world exists. Viewing the world in this way is viewing it sub specie aeternitatis — from the point of view of eternity — removed from it and disinterested” (2008: 326).

In his doctoral thesis ‘Tractarian Mysticism’ (1999) Woodruff explains the sub specie aeternitatis gaze as: “contemplating an object in the world as a world, as a limited whole. It also means willing that it be just the object-world it is and no other. To attend to an object in this light – with this attitude – is to see it as beautiful. ‘And,’ Wittgenstein says, ‘the beautiful is what makes happy’ (NB 21.10.16). [...] To see the world – under eternity’s aspect – is to ‘see’ its value, meaning, or significance” (1999: 128). Thus, what ethics and aesthetics have in common is their “gaze”, the particular way to look at

their object of concern. In aesthetic contemplation, “attention to an object [allows us to see the object] as a being purely for itself. [...] Similarly, [Wittgenstein] takes ethical contemplation as attention to the world as a being purely for itself, without a history or a future” (1999: 88). In short, looking at something *sub specie aeternitatis* is “[t]o attend to an object or a world in the eternal present, that is, under eternity’s aspect, is to contemplate it as an autotelic being” (ibid.).

6.44 Not how the world is, is the mystical, but that it is. What the artist delights in is not the facts being one particular way rather than another, but there being any facts at all — in effect, there being the possibilities there are.” (326)

Looking at the world as it is, without wishing to change it, nor viewing only what we think is valuable (remember that, for Wittgenstein, neither can we truly claim that we can cause anything in the world, nor that we can make any meaningful claims about value in the world) is to look at it *artistically*; that is, just like the artist looks at objects and sees them exactly as they are. Moreover, this gaze makes the artist happy, since it reveals the beauty of the object she is viewing, allowing her to contemplate such beauty. A beauty that has nothing to do with her, but only with the object as it is. What is “good” – the ethical claim – thus, is merely that there are things in the world that are in some way:

“It now seems that this is precisely what Wittgenstein means to say. The things which have real value are not facts, not anything in the world. What has real value are just: a certain kind of outlook, combined with the action of adopting that outlook; and what that outlook finds in looking at the world. What is really good, it seems — what the artist’s eye delights in — is not the facts being any particular way, but simply their being some way.” (Morris 2008: 328).

(5.61) What we cannot think, that we cannot think: we cannot therefore say what we cannot think.

Now, the problem is that whatever it is that we can see through the ethical/aesthetic gaze, we cannot speak of. This is because the business of philosophy is that of saying “what is necessary, or possible, or impossible, *in the world*” (2008: 328, my italics); thus, nothing philosophical can be said about value. On the other hand,

6.522 There is indeed the inexpressible. This shows itself; it is the mystical.

Since Wittgenstein recommends that we adopt a mystical attitude – that is, the “aesthetic” attitude – as opposed to a philosophical one, then philosophy and mysticism must be two different attitudes:

“The difficulty with philosophy is that it attempts to say what cannot be said. But this is not what mysticism does: its attention is directed to that the world is (6.44), that there is anything at all, not to how the world is (which is what can be said). It does not attempt to say anything: it is a view (‘*Anschauung*’, 6.45) of the world, a way of feeling the world (‘*Gefühl*’, 6.45)” (2008: 337).

The problem is that there seems to be no place at all for philosophy in Wittgenstein’s work, since philosophy poses questions that concern what lies outside of the world (e.g. questions about value) and attempts to provide answers that cannot but be meaningless (6.52). In other words, philosophy has the ambition to speak of what cannot be said; thus, it seems that the only way out is to replace it entirely with the mystical attitude.

This, on the other hand, raises a paradox: did Wittgenstein write a philosophical treatise only to say that philosophy is nonsense? If so, isn't then the whole *Tractatus* – *qua* philosophical – nonsense? Wittgenstein is aware of this paradox, and replies that:

(6.54) My propositions are elucidatory in this way: he who understands me finally recognizes them as senseless [unsinnig: nonsense], when he has climbed out through them, on them, over them. (He must so to speak throw away the ladder, after he has climbed up on it.) He must surmount these propositions; then he sees the world rightly.

Here the ladder is the symbol for a particular kind of ascension: that of the reader who has followed Wittgenstein all the steps through, until she can clearly see this paradox, the fact that all the steps she has taken were, actually, nonsense, but the point she has reached is, on the other hand, meaningful. In his letter to the publisher of TLP, Wittgenstein wrote:

“[T]he book's point is ethical. I once meant to include in the preface a sentence which is not in fact there now, but which I will write out for you here, because it will be a key for you. What I meant to write then was this: my work consists of two parts: the one presented here plus all I have not written. And it is precisely this second part that is the important one. For the ethical gets its limit drawn from the inside, as it were, by my book; and I am convinced that this is the ONLY rigorous way of drawing that limit. In short, I believe that where many others today are just gassing, I have managed in my book to put everything firmly into place by being silent about it.”

(WSP: 94–95)

The ladder offered by TLP is one through which the reader can learn the ethical/aesthetical attitude in order to see what it is that Wittgenstein has been silent about; that is, in order to see the world rightly. In short, Wittgenstein has not tried to communicate truths, since any attempt would have ended up in nonsense; rather, he has attempted to *show* truths (if there are any). The reader can only view such truths if she adopts the mystical outlook, only after she has climbed on top of a ladder made of nonsense prepositions, which she is required to recognize as such and, thus, to throw away. The reader is required to adopt the mystical outlook without accepting the TLP as true, because only by doing so would she be able to perceive not much reality itself, but the *form* of reality.³⁹ The form of reality is fully appreciated when the reader has become an artist, when she can finally experience the wonder of seeing what there is as it is, and does not feel compelled to force it into propositions:

7 Whereof one cannot speak, thereof one must be silent.

Wittgenstein's Platonic idea of moral transformation through ascension and contemplation, as well as Wittgenstein's equation of ethics and aesthetics, were further developed by Iris Murdoch.⁴⁰ In particular, she develops her account of moral action around the concept of "loving attention", which is what allows one to contemplate an

³⁹ This method is what Woodruff (1999) calls "Trans-logical" (from Pseudo-Dionysius): "Being outside every negation marks it as incommensurable with a world whose possibilities dissolve at the limits of logic" (109). This is also expressed in TLP 6.522, where "[w]e cannot say in what way 'the sense of life' consists; there is no 'positive way', for Wittgenstein. Yet by reliance on a language whose logic is the world's 'great mirror' we can say in what this sense does not consist; this is the 'negative way'" (1999: 112). Only by being silent about it, could Wittgenstein "put everything firmly in place" (ibid.).

⁴⁰ For more on how Murdoch was inspired by Wittgenstein, see Caprioglio Panizza 2022.

other's reality as it is and, thus, gain access to truth. In this picture, radical moral change takes the shape of an ascension insofar as the agent stops being the centre of attention and, thus, the individual gaze transcends the individual. This, according to Murdoch, is what radically transforms agents. It is, therefore, to Murdoch that I now turn.

2.1.3 Moral Transformation in Murdoch

Murdoch holds a pessimistic view of moral change: according to her, Freud was right in seeing the “psyche as an egocentric system of quasi-mechanical energy, largely determined by its own individual history, whose natural attachments are sexual, ambiguous, and hard for the subject to understand or control” (OGG: 51/341).⁴¹ If this is indeed the case then, *pace* existentialists, we are not that free when it comes to choosing to act in one way or another, when we impose our will on the world: “The area of [the moral agent's] vaunted freedom of choice is not usually very great” (SGC: 78/364). We are naturally drawn away from reality by our egocentric drives;⁴² we are beings whose centre of concern revolves around themselves. Doing something about it is, thus, incredibly difficult, since “fantasy is a stronger force than reason. Objectivity and unselfishness are not natural to human beings” (OGG: 51/241). Thus, it is moral philosophy's concern to find ways for agents to resist these natural drives, such as that of self-centered fantasy (OGG 344). Although difficult, it is possible: Murdoch endorses

⁴¹ According to Murdoch Freud provides “a realistic and detailed picture of the fallen man” (OGG: 51/241), where the ego is constant source of distraction from what is truly there. This is why “continuous control is required” (MGM 269).

⁴² See also MGM 296: “Our confused conscious being is both here and elsewhere, living at different levels and in different modes of cognition. We are ‘distracted’ creatures, extended, layered, pulled apart.”

what has been called Platonic moral realism, according to which there is such a thing as reality, which is to be found *outside* of the self but still, contra Wittgenstein, *in* the world.⁴³ What can trigger moral transformation, for her, is the process of ‘unselfing’: that is, of paying attention to what is other than one’s self.⁴⁴

Unselfing is the process whereby “[w]e cease to be in order to attend to the existence of something else” (OGG: 59/349). In ‘The Sublime and the Good’ Murdoch explains unselfing by drawing an analogy with art. According to her, both art and moral action require love, where love is the discovery that there is something other and independent of my own self:

“Art and morals are [...] one. [...] The essence of both of them is love. Love is the perception of individuals. Love is the extremely difficult realisation that something other than oneself is real. Love, and so art and morals, is the discovery of reality.” (SBL 225)

As it was for Wittgenstein, here again we can see that art and morals consist of the same activity; they require the same attitude. For Wittgenstein, it was contemplation of the form of reality, for Murdoch, the activity of both art and morals is paying loving attention. Seeing something other than one’s self as real and independent *just is* moral

⁴³ This is the point of departure with Plato (and Aristotle): virtue is not good for the agent; it is rather pointless (OGG: 71/358; SGC: 78/364). There is no telos in human life (SGC: 79/364).

⁴⁴ One can love, and direct loving attention to, another whose deficiencies and faults she fully recognizes (Wolf 2014). Cordner (2016) and Wolf (2014) emphasize that it is the person as a whole that is the proper object of loving attention.

(and artistic) action.⁴⁵ In this picture, fantasy is the worst of the enemies of both art and morals, since it does not allow us to see others as they are, but, rather, as we wish them to be, or according to our dreams and fantasies, which are peculiarly our own; again, the ego that speaks instead of reality. For this reason, “Fantasy, the enemy of art, is the enemy of true imagination: Love, an exercise of the imagination.” (SBL 225).

Fantasy and imagination, thus, are two different capacities. As mentioned before, fantasy is the projection of one’s ego to others, whereby one does not truly see others’ reality, but merely a version of one’s own ego-driven thoughts and concerns. Imagination, on the other hand, is the capacity to conceive of others’ independent existence, to entertain all the possible modes of existence that are not our own. It is what allows us to get out of our own egocentric thoughts and imagine what it is to be something other than ourselves. Imagination is also what allows the good writer to build characters that are not herself, or a projection of her ego.⁴⁶ In particular, writers of tragedies are the ones who are the most imaginative, since they truly need to be able to perceive characters’ beings, with their intricacies and their peculiarly human dynamics, without interfering. It is in this sense that being able to love is somehow tragic:

⁴⁵ See Caprioglio Panizza (2022) for a comprehensive account of Murdochian attention as “a truth-seeking engagement of the individual with reality” (24), where the “truth-seeking” aspect is what makes attention a *moral* capacity, in its allowing us to see the truth and act on it. It is in this sense that if we wish to be ‘good’ moral agents, we should be attentive moral agents.

⁴⁶ “In the case of Tolstoy and Pasternak, it is, I think, not difficult to see that the quality of their greatness should be called compassion, love: the non-violent apprehension of difference. And with what exhilaration do we experience the absense of self in the work of Tolstoy, in the work of Shakespeare. That is the true sublime” (SBL 227).

“The tragic freedom implied by love is this: that we all have an indefinitely extended capacity to imagine the being of others. Tragic, because there is no prefabricated harmony, and others are, to an extent we never cease discovering, different from ourselves. Nor is there any social totality within which we can come to comprehend differences as placed and reconciled. We have only a segment of the circle. Freedom is exercised in the confrontation by each other, in the context of an infinitely extensible work of imaginative understanding, of two irreducibly dissimilar individuals. Love is the imaginative recognition of, that is respect for, this otherness” (SBL 226)

The analogy with literature clarifies also the analogy between the artist and the moral agent.⁴⁷ It is not the case that artists do not create anything, but rather just contemplate objects in the world that already exist – this would not be credible. Just like the writer creates her characters as beings independent of herself, the artist created the work of art with the same love and respect for its being:

“In the creation of a work of art the artist is going through the exercise of attending to something quite particular other than himself. The intensity of this exercise itself gives to the work of art its special independence. That is, it is an independence and uniqueness which is essentially the same as that conferred upon, or rather discovered in, another human being whom we love” (SBL 228).

Moreover, art, just like good character building and moral action, “is a thing totally opposed to selfish obsession” (SGC 371). Not only it is guided by a concern for reality, but also, in Platonic language, “inspires love in the highest part of the soul” (ibid); that kind of love which is “unpossessive”, self-less contemplation and appreciation.

⁴⁷ This is one of Murdoch’s points of departure from Plato (SGC 373).

The fact that seeing, or attending to, others lovingly *just is* moral action needs more clarification. There are two elements to expand on in order to understand this claim: Murdoch's Platonic moral realism assumes that agents perceive the authority of the Good; that is, that agents – sometimes – succeed in seeing reality because they are attracted to the Good. The Good has some sort of magnetic force (MGM 14; IP 336) that allows one to redirect one's gaze outside of her egocentric concerns.⁴⁸ The second element to consider is that, when one manages to finally see reality as independent of one's self, then it is hard to ignore what one has seen: that is, that there exist such things as other beings. It is in this sense that, when one manages to see something lovingly, then one cannot but act in accordance with what she sees;⁴⁹ in Murdoch's words, it is in this sense that “true vision occasions right conduct” (OGG 355). She goes on to say that “[t]he more the separateness and differentness of other people is realised, and the fact seen that another man has needs and wishes as demanding as one's own, the harder it becomes to treat a person as a thing.” (ibid.).⁵⁰ In what follows I will expand on Murdoch's Platonic realism.

⁴⁸ This is also present in her account of Perfectionism, where the idea of perfection produces “an increasing sense of direction” (OGG 355).

⁴⁹ See also Caprioglio Panizza on “moral affordances” (2022), more on this also in chapter III.

⁵⁰ It is important to note that when Murdoch speaks of “others”, she does not only consider people, but also natural entities, animals and objects (White 2020): hence, the famous quote: “I am looking out of my window in an anxious and resentful state of mind, oblivious of my surroundings, brooding perhaps on some damage done to my prestige. Then suddenly I observe a hovering kestrel. In a moment everything is altered. The brooding self with its hurt vanity has disappeared. There is nothing now but kestrel” (SGC in ... 370).

First of all, note that another Wittgensteinian element that we can find in Murdoch is the importance of *forms*.⁵¹ Wittgenstein has – at least according to ethical interpretations – built the TLP so that the reader could reach the point of perceiving the form of the truth, rather than its content. Similarly, Murdoch’s Platonism conceives of moral knowledge as that which we arrive at through unselfing, and which appears to us not much as the content of the Good, or as its incarnation in something wordly, but, rather, as the *form* of the Good. Through loving the good in particular things, we ascend “to an understanding of Good itself” (Bloom 2023). Also in Murdoch, thus, we can find the theme of ascension. In SGC she employs Plato’s allegory of the Sun in the myth of the Cave to explain her own conception of the Good, and human ascension to it (SGC 376).

“The sun is seen at the end of a long quest which involves a *reorientation* (the prisoners have to turn round) and an *ascent*. It is real, it is out there, but very distant. It gives light and energy and enables us to know the truth. In its light we see the things of the world in their true relationships” (SGC 376, my italics). That is, we see what really matters, what really there is to know, independently of what we think it should be, of our own egotistic concerns. The reorientation is the human gaze that is now focused on reality; this also coincides with the ascent. She goes on to explain what it means to see things in the sun’s light:

⁵¹ Bloom (2023) notes that Murdoch does not subscribe to the aspect of Plato’s view that regards the forms as more real than individual objects and persons who partake of them in the world of experience, nor as inhabiting a transcendent world beyond our world of experience (Hämäläinen 2019: 267). As a matter of fact, she seems to reject this as a proper interpretation of Plato (SGC: 96/378f; however, see Robjant 2012 for the opposite view).

“Looking at it itself is supremely difficult and is unlike looking at things in its light. It is a different kind of thing from what it illuminates. The proper and serious use of the term refers us to a perfection which is perhaps never exemplified in the world we know [...] and which carries with it the ideas of hierarchy and transcendence” (ibid).

Thus, seeing reality under the light of the Good does not imply seeing the Good: “We see differences, we sense directions, and we know that the Good is still somewhere beyond” (ibid.). The quest is infinite because on the one hand we perceive the magnetic force of the Good and, on the other, we cannot directly access it; that is, perception of the Good in the world is always perfectionable. Murdoch rejects common ways to make sense of the term “good”, that understand it as concepts such as “functional” (e.g. ‘the good knife’), or “well-behaved” (e.g. ‘the good fellow’). In her words, good “is not a mere value tag of the choosing will” (SGC 377). It is something that will always eschews our conceptualizations, yet it will always guide us towards an ascension.⁵²

In IP Murdoch provides a fictional example of moral action that is guided by loving attention, with the aim of clarifying what ‘unselfing’ really amounts to, given our natural self-centered tendencies:

A mother, whom I shall call M, feels hostility to her daughter-in-law, whom I shall call D. M finds D quite a good-hearted girl, but while not exactly common yet certainly unpolished and lacking in dignity and refinement. D is inclined to be pert and familiar, insufficiently ceremonious, brusque, sometimes positively rude, always tiresomely juvenile. M does not like D’s accent or the way D dresses. M feels that her son has married beneath him. Let us assume for purposes of the example that the mother, who is

⁵² As Murdoch nicely puts it, “[a] genuine mysteriousness attaches to the idea of goodness and the Good” (SGC 381).

a very ‘correct’ person, behaves beautifully to the girl throughout, not allowing her real opinion to appear in any way. We might underline this aspect of the example by supposing that the young couple have emigrated or that D is now dead: the point being to ensure that whatever is in question as *happening* happens entirely in M’s mind. Thus much for M’s first thoughts about D. Time passes, and it could be that M settles down with a hardened sense of grievance and a fixed picture of D, imprisoned (if I may use a question-begging word) by the cliché: my poor son has married a silly vulgar girl. However, the M of the example is an intelligent and well-intentioned person, capable of self-criticism, capable of giving careful and just attention to an object which confronts her. M tells herself: ‘I am old-fashioned and conventional. I may be prejudiced and narrow-minded. I may be snobbish. I am certainly jealous. Let me look again.’ Here I assume that M observes D or at least reflects deliberately about D, until gradually her vision of D alters. If we take D to be now absent or dead this can make it clear that the change is not in D’s behaviour but in M’s mind. D is discovered to be not vulgar but refreshingly simple, not undignified but spontaneous, not noisy but gay, not tiresomely juvenile but delightfully youthful, and so on. And as I say, *ex hypothesi*, M’s outward behaviour, beautiful from the start, in no way alters (IP: 329).

In this famous example M’s redirected gaze allows her to *truly* see D. But what is it that M actually *does*? In which sense is it moral *action*? The short answer is: she puts herself aside for a moment. Murdoch works with a conception of the self that is *substantial* – i.e. there is such a thing as a self that can be improved – and that finds some degree of unity in the moral self. Thus, the action that underlies ‘unselfing’ cannot be that of annihilating the self. According to Murdoch, there is only a part of the self that is problematic, which is that of the ego; that is, the part that urges us to focus on our self-concerns, and that prevents us from looking at the world with a gaze that is independent

of such concerns.⁵³ In this sense, M's putting herself aside is understood as M re-directing her gaze from her own self concerns to D; loving D by seeing D as a reality that is independent of M's egocentric concerns. Murdochian moral realism is exemplified by the fact that M really does something morally *better*, as opposed to what she used to do before, where "do" indicates action in Murdochian terms, that is inner, invisible, but morally relevant action.

Finally, she uses the Platonic metaphor of vision exemplified, once again, by the myth of the Cave, to explain precisely in what way it is that M is doing something morally better than before:

"I suggest we can if we simply introduce into the picture the idea of attention, or looking [...]. I can only choose within the world I can *see*, in the moral sense of 'see' which implies that clear vision is a result of moral imagination and moral effort. There is also of course 'distorted vision', and the word 'reality' here inevitably appears as a normative word. When M is just and loving she sees D as she really is. One is often compelled almost automatically by what one can see" (IP 1999: 331-332).

Just as those who escape the Cave can finally see (but not stare at) the sun and its rays that finally shed light on the real world, attentive agents can finally have a glimpse of the Good and its illuminating rays on reality.⁵⁴

We have seen that when it comes to Moral Transformation, the dynamics of transformation in the moral domain are likened to those in the aesthetic domain. In Plato,

⁵³ See Caprioglio Panizza (2022), where she calls Murdoch's account of the self as "the Tame View", since in order to morally ascend we do not need to eliminate the self, but to suppress, or manage, some parts of it.

⁵⁴ Because "[t]he Good itself if not visible" (Murdoch OGG 1999: 358).

Wittgenstein and Murdoch, transformation involves ascending towards reality, even if we cannot fully reach it. Aesthetics is the attitude required in order to transcend and, thus, be morally transformed, since it is aesthetic contemplation that allows moral agents to see reality in its forms, independently of the agent. Thus, if we must draw some preliminary conclusion on the conceptual analysis of moral transformation, we can say that the process of ‘going beyond’ requires, in the moral case, to transcend one’s self-centered drives and contemplate what there is to be seen. This is both aesthetic and moral action, if we assume, with Murdoch, that reality exists and it has a magnetic force. Contemplation is a moral activity which is inherently transformative. To end this section, I will consider a more current account of “transformation as an activity”, which is that of Agnes Callard, and stress its many points of departure from Murdoch, as well as with the aesthetical/ethical understanding of Moral Transformation that I have provided so far.

2.2 Transformative Activities

For Wittgenstein mysticism is an activity; it is ascending to the place where truth can only be perceived through what it cannot be said about it, through its contours. This ascension is an ethical transformation, since the endpoint is the perception (the view) of the form of values. For Murdoch looking at something with the loving gaze is itself a moral action. In her view moral transformation is explained through Plato’s allegory of ascension, but it is less obviously an ascension per se. The direction it takes is towards the Good and since in Plato’s allegorical imagery the Good is somewhere above us, then it takes the form of an ascension – but it could as easily be a descent towards something deep, or an oblique trajectory towards what eludes the self-absorbed eye. What matters

is that paying loving attention to what is other than one's self is already a moral action, and one that is transformative insofar as it allows one to access reality, and to be guided by it. It is transformative, because Murdoch's assumption is that in everyday life, we are slaves of our own egotistic forces, so we see the world through our self-absorbed lenses. Thus, what we gain access to through loving attention is likely to be something we have not seen yet.

Now, the activity in both Wittgenstein and Murdoch is pretty passive: it involves looking, contemplating, a moment of letting go (Ollson 2018). More recently, however, moral transformation has been associated with activities which are much more obviously active than merely standing by and contemplating objects in the world. Today, "transformative" is widely associated with what Murdoch would call acts of the will, such as choices (e.g. becoming a parent, Laurie Paul 2014) or highly engaging experiences (e.g. playing games, Nguyen 2019), which also have a moral counterpart in their having an impact on the agent's action-guiding values, which are likely to change more or less drastically after the choice or the activity is brought about. Thus, the interpretation of moral transformation today goes in a diametrically opposed direction: it is not passive contemplation, accompanied by the silencing of one's concerns, but rather an active engagement with one's self concerns.

It is important to stress that moral transformation needs to be differentiated from what Laurie Paul named "transformative choices" (2014). Paul's inquiry belongs to the field of action theory and, thus, it is neither concerned with the mechanics of self-transformation, nor with the peculiarly moral element that is present in some self-transformations. Drawing from Ullman-Margalit (2006), Paul explores the – alleged – irrationality of those kinds of choices that imply a personal transformation, where one has to decide on preferences that will likely not be present anymore once the transformation has taken place. The paradigmatic examples she picks to exemplify such

transformative choices are becoming a parent and becoming a vampire. On what rational grounds will I base my choice to become a parent or a vampire, if once I have become a parent and a vampire I will be basing my choices on completely different grounds? The answer need not concern us here; some philosophers are unconvinced by Paul's reading of transformative choices and their alleged decisional irrationality, nor – crucially – of their alleged *moral* element.⁵⁵ Among those, I will consider Agnes Callard, who draws an interesting distinction between Paul's transformative choices and what she labels “transformative activities” (2020), which seems to go back in the direction of moral transformation.

Callard considers Elena Ferrante's series of novels *My Brilliant Friend*: the friendship between Lila and Lenù is experienced by the two friends as a continuous competitive activity that is transformative in so far as they each aspire to become the other.⁵⁶ Differently from the cases considered by Ullman-Margalit and Paul, where the agent undergoes a breakage between her self-before-the-choice and her self-after-the-choice, in transformative activity “the agent actively works to become a new kind of person

⁵⁵ For example, Kristjánsson stresses that a transformative choice has nothing to do with what he labeled “epiphanic moral conversions”, since “[w]hile [Laurie] Paul's account opens the door to self- elicited transformations, flowing from deliberate decisions, those are not tantamount to decisions to undergo a moral conversion; she is rather describing the possibility (and commonness) of decisions to take spectacular jumps into the unknown— and those do not necessarily alter the orientation of one's moral compass.” (Kristjánsson 2018: 25).

⁵⁶ According to Callard's aspiration model (2018), when one decides to transform ones' self, proleptic reasons. She talks of transformations in general, but it can be applied also to moral transformations. “Aspirants try to get some new values into view” (2017: 1). Self-transformations are “both active and transformative” (2017: 5): counter-vegetarianism on her visit to Hungary. Aspiration: like Callard becoming a classicist (different from self-transformations). In self-transformation the agent already knows what she will gain; in aspiration the agent does not entirely know what she is getting into, it is more tentative (2017: 8).

without undergoing a break in her biography” (2020: 149). In particular, in “transformative revelations” such as those depicted by Ullman-Margalit and Paul, “you are transformed by what happens to you, whereas in transformative activities, you yourself do the transformation”⁵⁷ (151). Finally, transformative activities constitute a particular kind of learning experiences; that is, one in which every part of the transformation is a learning activity, one of “inquiry.” Knowledge does not “arrive” *after* the transformation as it is the case for transformative revelations; in transformative activities the agent figures it out. Let us go back to the competitive relationship present in *My Brilliant Friend*, which Callard calls “aspirational competition” (157). As she notes, Lila and Lenù, the two protagonists of the novel, grow up in a poor post-war neighborhood of Naples, where their options – especially as women – are limited. The doom of living their mothers’ lives is present throughout the entire book series. As a child, Lenù senses her own fear of becoming like the adult women around her, but she does not know any other way; when she gets to know Lila, she sees something special in her, something that may help her see other options, other values, and allow her to escape her fate:

“...something convinced me, then, that if I kept up with [Lila], at her pace, my mother’s limp, which had entered into my brain and wouldn’t come out, would stop threatening me. I decided that I had to model myself on that girl, never let her out of my sight, even if she got annoyed and chased me away” (Ferrante, p.46 in Callard 2020: 159).

At the same time, through Lenù’s first-person narration, we get to see the tension in their relationship, since Lila seems to her the truly brilliant one, the girl who is too poor

⁵⁷ See also Ruth Chang (2015) on event-based and choice-based transformations.

to pursue her education, but who gets to learn ancient Greek and even ends up helping Lenù with her homework translations. Through Lila, Lenù gets to learn what she would like to become – she aspires to be Lila. At the same time, if we detach a little from Lenù's perspective, we can see that also Lila experiences something similar. It is because she wishes to prove herself to Lenù that she is constantly driven to education and, later on, to politics and literature. We do not have access to her inner world, but we can imagine that she also aspires to be Lenù, the girl who goes to school and, later on, the feminist writer. In short:

“Together, they begin to see the alternative possibilities—alternative values, activities, ways of being—that constitute, for each girl, a way of life. Her relationship with Lila strikes Elena as a game which allows them to escape the strictures of their environment. [...] The goal of being and beating Lila is the closest Elena [Lenù] can get to the idea of the value of escape, just as getting a higher grade on a test can be the closest that you get to the idea of mathematical excellence” (160).

Thus, Callard concludes, transformative activity is “a way of learning to be different by acting differently” (ibid.), where while you act you do not fully know what it is to be that way. The competition, in this case, is not one in which the agents want to win something over the other, but one in which one wishes to become the other.⁵⁸

Callard's account of transformative activity can include activities that are peculiarly moral: one can aspire to become compassionate, inclusive, courageous, or any virtue we can think of, and this aspiration can take the shape of a learn by doing. After all, she reminds us, Aristotle himself claimed that learn by doing is the path towards virtue,

⁵⁸ This seems to be supported by the Christian idea of alter-Christuts, where aspiring to sanctity involves aspiring to become Christ-like. On the role of imitation Christi in Aquina's moral exemplarism, see Vaccarezza (2018).

since it is what grants the re-shaping of habits into virtuous habits. Significantly, one can aspire to become virtuous by competing with already virtuous people in a way that is similar to that of Lenù with Lila; that is, by allowing the other to disclose virtue in a way that the not-yet-virtuous person could not even conceive of. The bitterness experienced when in the presence of the virtuous agent need not lead to (self-)destructive envy or depression; just like with Lenù, it may trigger the competition to become that person and be better than them at being them.

Interestingly, in SBL Murdoch warns us against understanding ascension, both moral and aesthetic, as if morals and art were pedagogical devices:

“To say that the essence of art is love is not to say, is nothing to do with saying, that art is didactic or educational. It is of course a fact that if art is love then art improves us morally, but this is, as it were, accidental. The level at which that love works which is art is deeper than the level at which we deliberate concerning improvement. And indeed it is of the nature of love to be something deeper than our conscious and more simply social morality, and to be sometimes destructive of it” (SBL 227).

This seems to suggest that Murdoch distinguishes between improvements that are triggered accidentally and those that are decided beforehand. Moral ascension through love seems to be of the first sort. Yet she too employs the example of learning a language to further explain her account of loving attention:

“The τέχνη [science, craft] which Plato thought was most important was mathematics, because it was most rigorous and abstract. I shall take an example of a τέχνη more congenial to myself: learning a language. If I am learning, for instance, Russian, I am

confronted by an authoritative structure which commands my respect. The task is difficult and the goal is distant and perhaps never entirely attainable. My work is a progressive revelation of something which exists independently of me. Attention is rewarded by a knowledge of reality. Love of Russian leads me away from myself towards something alien to me, something which my consciousness cannot take over, swallow up, deny or make unreal. The honesty and humility required of the student - not to pretend to know what one does not know - is the preparation for the honesty and humility of the scholar who does not even feel tempted to suppress the fact which damns his theory” (IP 373 – 374).

Here, it seems that a learning activity may also provide the agent with the possibility of engaging lovingly with it and, thus, to experience ascension. Thus, there is a sense in which by engaging in some activity that is not itself moral, we can learn the loving attitude that is necessary for moral action. If we add the aspiration element, then we can see that I could, in principle, engage lovingly in an activity I do not love yet, with the aspiration of coming to love it. More generally, I can aspire to develop the loving gaze and, in Callard’s terms, I can learn to love by loving. In this case, I can learn to love an other’s reality by learning a language. There is an element of instrumentality that is alien to Murdoch: for her I do not learn Russian to learn how to love something other than myself. I just learn Russian, and I truly learn Russian if I have a loving attitude towards it. Moreover, there is no competitive element in Murdoch’s account. As a matter of fact, that of Lila and Lenù is exactly the opposite of the M&D example.

To begin with, Lila is always, inescapably seen through Lenù eyes: never does the reader have the chance to truly see Lila independently of Lenù’s anxious concerns about herself and how she appears to others. To be clear, in my view this is not a fault of the book series: quite the contrary, it is one of its most interesting elements, that of the total

absorption of Lenù in her own self-concerns – not to mention that Murdoch's novels are quite similar in this respect, since what she wishes to depict is precisely the fact that people are naturally self-absorbed and only see others as they *think* they are, not as they truly are. Then, competition itself, even if it takes place among individuals, is a self-directed activity: it is a sign that one is looking at others through self-concerned lenses, since what one sees is what one would like to have.⁵⁹ The focus is still on the ego. Who the other person is apart from what she has, both in terms of achievements and character traits, is unimportant, and this cannot be less Murdochian. As mentioned before, when writing about learning Russian, Murdoch expresses what it is that love *for* Russian does, not the aspiration to become a person who knows Russian does. Love for Russian allows Russian to reveal itself as it is; the gaze is oriented towards Russian. On the other hand, in aspiring to become someone who knows Russian, the gaze is oriented towards Russian only instrumentally, to achieve some personal goal. This may lead to more or less drastic changes in one's personality, but does not guide towards ascension. Similarly, Lenù sees Lila only instrumentally, as a way to become her. Yes, this may lead Lenù to become better in some respects – more literate, more outspoken – but does not lead her to truly see Lila, who remains a mystery throughout the book series.

We can see that the aesthetic/ethical reading of Moral Transformation has a necessary element of selflessness. On the other hand, the current account of transformations that – arguably – have moral significance (transformative experiences, transformative activities) go in the opposite direction: the focus is on the self through and through; on her concerns, on her goals, on her wishes to change, on what she thinks she needs in order to become like someone else. There is no room for what there is apart from the self.

⁵⁹ See Vaccarezza and Niccoli 2022 on the power of eliciting envy.

We can conclude that if the elements of moral conversion were those of radicality and of experienced abruptness, moral transformation on the other hand can be much more “linear;” that is, moral change can come about, or be experienced as the gradual result of a journey, be it an ascension journey, or an activity journey. This does not rule out the possibility of an “epiphanic” experience of moral transformation, especially when it comes to ascension-like transformations, since what is slowly gained is contact with reality, and disclosure of moral reality, as we have seen, is likely to trigger moral conversions. Thus, we can already spot some possible overlap between moral conversion and moral transformation, as well as at least one fundamental point of departure: “transformation” describes something gradual, while “conversion” describes something abrupt. I will now turn to the final part of this chapter, which is that on “moral progress.”

3. Moral Progress

Broadly, “moral progress” indicates changes for the better that can happen at the level of institutions (structural level) and/or the level of individuals.⁶⁰ The very fact that I can write this dissertation indicates that there has been some level of societal progress concerning the role of women in society. “Progress” implies a move forward; in this case, a move away from patriarchal towards a more egalitarian society. “Moral” implies that this move concerns changes in attitudes, behaviors and value concepts that can be evaluated. In this case, we could say that the fact that now women have more options than they used to is a *good* state of affairs; indeed, a *better* one compared to before. This is a description: there has been some moral progress. One may also wish to look for the dynamics of societal moral progress, how it comes about, when it is desirable to trigger and whether we can give it a direction. These are widely discussed issues that have yet to be settled.

When it comes to individuals, the issue is far from being settled either. To begin with, one may even wonder whether it is a thing in itself, and not merely a reverberation of societal moral progress. After all, societal moral progress has an undeniable impact on individuals: at some point, racism, sexism and homophobia have been framed as “problems”. This implies that racists, sexist and homophobic individuals, when openly discriminating, do not perform actions that are “acceptable” anymore, but that are framed in a morally charged way as “wrong”. People, thus, may progress just in virtue of these massive conceptual changes at the level of their culture. On the other hand, individuals have an undeniable impact on structural moral progress (Moody-Adams

⁶⁰ See also the concept of moral progress suggested by Jamieson 2002, where moral progress occurs when “a subsequent state of affairs is [morally] better than a preceding one, or when right acts become increasingly prevalent” (20).

1999; they reciprocally influence each other, Buchanan, Powell 2018, 210-11; Kumar, Campbell 2022, 135-138). Structures do not change unless people force them to, or advocate for change, sometimes at the cost of their lives. Individuals are, thus, both the engines and the recipients of structural moral progress.⁶¹

Now, if we are analytical about this discussion, and if we think of human beings as beings that, a bit like cyborgs, follow instructions rationally, who can “update” their knowledge successfully and without any distress, then we may reach the following conclusions: after racism has been framed as “wrong”, then all forms of racism ceased to exist; after sexism has been framed as “wrong”, then all forms of sexism ceased to exist; and so on. But this is obviously not true. In being recipients of structural moral progress, individuals are particularly *resistant* to it. It is obvious, then, that there are some peculiar features of being an individual agent that do not belong to structures, that are crucial for moral progress and that are, thus, worth considering (Archer 1995).⁶² To these peculiarities scholars have turned in order to understand what it is *for an agent* to morally progress (Buchanan, Powell 2018; Kumar, Campbell 2022; Schinkel & De Ruyter 2017; Severini 2021; Songhorian et al. 2022), and if it is at all possible (some say it is not, see Klenk & Sauer 2021). For the purpose of this chapter, let us turn to what we can call “individual moral progress”.

3.1 Individual Moral Progress

Individual Moral Progress generally indicates changes for the better that have individuals as subjects. “For the better” can express an implementation in *moral capacities*, that is, in people’s “exercise of human moral powers – their capacities for

⁶¹ See Madva 2016 for a dissenting view.

⁶² Although some scholars do not agree. See Buchanan & Powell 2018, Macklin 1977, Sauer 2023.

having moral concepts, making and appreciating moral arguments, being committed to moral consistency, and having moral motivations” (Buchanan & Powell 2018: 50). “For the better” can also imply that some form of change has taken place and it is better from a moral point of view, but it need not require an improvement in morality itself (Sauer 2024). In the former case, the progress is understood as “narrow”, while in the latter it is understood as “broad”; as in, narrowly having to do with morality being the very thing that progresses, and broadly as having to do with changes that are also morally good (Klenk & Sauer 2021, Sauer 2024).⁶³

There is a vibrant debate about the individual dimension of moral progress. On the one hand, some scholars identify cognitive processes as the engine of moral progress, where moral progress comes about through changes in conceptualization (Moody-Adams 1999; Hermann 2019; or in “understanding”, Severini 2021). On the other hand, Rorty 1999 defends an “affective view” according to which changes for the better can only truly come about through emotional drives.⁶⁴ Consider the concept of “sexual harassment”; as pointed out by Fricker (2017), before this term was introduced (1970s) women did not have the means to understand the nature of the wrongs that were done to them. Afterwards, behaviors that until then fell under the category of, say, “being effusive” could finally be understood as sexist, diminishing, disrespectful, etc. Concepts deepened (Platts 1988) so as to become more sophisticated, more nuanced. We can imagine men who, until then, believed that calling a female co-worker “darling” was perfectly fine and even respectful, deepening their understanding of situations and, thus, gaining the nuances needed in order to discern what is truly respectful and what is not. These men have, according to this highly intellectualist picture, morally progressed. We can also

⁶³ Sauer 2024 also explains the distinction in terms of the wrong and the bad: “Tsunamis are bad, but they are not wrong ... Wide moral progress, then, reduces the bad; narrow moral progress mitigates the wrong” (2024: 28-29).

⁶⁴ For a comprehensive overview see Klenk and Sauer 2022.

imagine them “updating” their moral knowledge on how to treat women through perfectly rational, non-biased, evidence-based and impartial cognitive processes.⁶⁵ The fact that sexual harassment is still widely present in working environments, as well as in ordinary daily activities such as going grocery shopping, reminds us once more that we are not cyborgs. If individual moral progress was due merely to un-biased cognitive processes, then individual moral progress would be impossible (Klenk & Sauer 2021). On the other hand, also purely emotivist accounts seem to miss on something. According to Rorty (1999) and Slote (2007), changes for the better in moral judgments can only be triggered by affective-emotional experiences. According to this picture, men would come to understand their sexist dispositions only if they have an emotionally impactful experience; say, their daughters come home one day and tell them that they have been victims of the kind of behavior that they deemed perfectly fine until a moment before. They see their daughters in tears, angry as they have never been before, telling them “why can’t they take me seriously? Why am I always degraded so much?”, or they hear their daughters’ grocery shopping horror stories, where men make comments, laugh creepily at what they say, follow them home, and so on. These men feel empathy, they experience all their daughters’ anger and finally understand what it means to be a woman. They swear they will never treat a woman like they used to. This may be more immediate than the kind of moral progress that happen only at the intellectual level, but it may still strike as way too arbitrary. There must be a moment in which agents, due to an emotional experience, reason about it in a way that includes not just what they have experienced, but their values scaffolding at large.

This is not the place to settle the dispute. However, what we can say is that an account of individual moral progress that is both true to human psychology and normatively

⁶⁵ For more on epistemic views of moral progress, see Kumar, Campbell 2022, 195-198; Singer 2005; 2011, Ch. 4–5.

feasible needs to propose a synthesis between the intellectualist and the emotivist insights. Here, I am interested in which are – if at all – the unique elements of the idea of individual moral progress; this is because we might just find out that that of “individual moral progress” is a redundant label, given all that we can already say about Moral Conversion and Moral Transformation. So far, the only element we possess is that, whatever the real dynamics of moral progress are – societal or individual, but also purely intellectual or emotivist – what it is that marks progress is the fact that some states of affairs are *better* than they used to be. This, however, was also a requirement of epiphanic moral conversions, which may not be a problem in itself, unless we find out that it is the *only* requirement for a process to be called “moral progress”. If it is indeed the case, then we should be eliminativist about the label “individual moral progress”.

However, we have good reasons to think that this is not the only necessary requirement for a process to be legitimately called “progress”. In fact, it may not apply at all: one may conceive of “progress” as a dynamic that is all but linear. In this respect, I will now introduce Wittgenstein’s own notion of progressive changes, which goes in a significantly different direction.

3.2 Wittgenstein on the illusion of progress

As mentioned before, “progress” means going forward, advancing. Imagine I have to reach the peak of a mountain; I start walking and after a couple of hours I stop for a snack break. I look behind me and I see the town I left far away in the distance. I look ahead and I see the peak of the mountain a bit more clearly, though it is still far and my sight cannot follow the path all the way up. I can say that I made some progress towards the peak of the mountain. After two more hours of hiking I stop and I cannot see the

town anymore, while I can clearly see the last 500m of height difference that I need to climb to reach the peak. I can say that I have made progress towards reaching the peak of the mountain; in fact, *more* progress if compared to my previous snack break. If we think of the dynamics of progress as a linear path towards reaching an objective, then it seems that we can judge, at least in principle, how far it is that we are from the objective, and how much of the path it is that we have already walked. Note that the analogy of the mountain hike is not exactly as the analogy of the ladder in Moral Transformation, at least in its Platonic reading. The difference is that we do not seem to know what there is to be seen from above the ladder; contemplation has no expectations on the object. Moreover, we do not really reach anything, since what guides us in the first place – the Good, Beauty – is not for us to be reached. We can only contemplate its form through improved, but not perfect eyes. On the other hand, progress has a much more pragmatic flavor: we see what it is to be reached – e.g. social equality, developing compassionate attitudes, etc. – and we set forth to reach it. At some point we can look back and say “well, we have progressed with respect to *that*, but there is still a lot to walk”. It is in this sense that all conceptions of individual moral progress that are currently available consider “progress” as a move headed towards an objective: being a better moral agent. They diverge when it comes to judging what makes a moral agent better than she used to be, and to how she can progress in this sense.

But is moral progress, both in its societal and in its individual declination, a thing at all? Do we really advance towards objectives, or is it just a nice fairy tale, an illusion?⁶⁶ As we noticed at the very beginning, people do change throughout their lives, at least in some respect, and some more radically than other. Moreover, throughout their lives people do change, or at least modify a little, their action-guiding principles and values. So we can legitimately say that there is such a thing as moral change, understood very

⁶⁶ See Sauer 2024 for a categorical reply: the question itself, whether X is progress, is itself regressive, par 1.2

broadly. However, the fact that there is such a thing as broadly construed moral change does not imply that a change that we can judge for the better is necessarily progressive, in its being a move away from something worse and better towards some objective. In short, we may wish to question whether we should think of progress in such a linear fashion.

Wittgenstein was skeptical of the notion of ‘progress’ understood as going forward. He considers two examples that could be read as instances of progress, but that in fact they are not, at least not intended as “going forward”. The first example is his own work as a philosopher, the second is philosophical work in general. As a matter of fact, PI is generally taken to be a move forward, a progress, in relation to his former work, the TLP. But Wittgenstein seems to warn against this reading:

“It is in the nature of progress that it is always less than it seems” (Nestroy).⁶⁷

According to Read (2016), this is to remind us that it is our tendency to “drastically, and dangerously, and in a dubiously self-congratulatory mode, to overestimate the progress that we have made as individuals, as a discipline, as a society, and as a civilisation.” (Read 2016: 267). The tendency in question is that of looking back at what we used to say and think that now we have much greater ideas that replace the old, rubbish ones; similarly, we have the tendency as a society to think that the past was dark and uncivilized, while now we have finally and brilliantly progressed from the ‘Dark Ages’; all this, according to Read’s reading of Wittgenstein, is to engage in “self-back-patting behaviour of a dubious kind” (2016: 268). Truth is, that we have not really advanced as much as it seems, or as we wished it to be, and this is true for both the work of

⁶⁷ Philosophical Investigations, epigraph, 1953.

philosophy and society at large. This, of course, does not imply that a move forward does not exist at all in either cases. Wittgenstein's point seems, rather, to be that we *overestimate* how much of a move forward what we call 'progress' actually is and, consequently, whether we should put all that emphasis on what we think of as instances of progress. Wittgenstein considers the scratching of an itching spot: "If somebody scratches the spot where he has an itch, do we have to see some progress? Isn't genuine scratching otherwise, or genuine itching itching? And can't this reaction to an irritation continue in the same way for a long time before a cure for the itching is discovered?" (Culture and Value: 86, in Read 2016: 268). The mere fact that the new state of affairs – absence of itching – is better than the former state of affairs – the presence of itching – does not imply that there has been progress, a going forward of some sort. It may just be a reaction, just like scratching where it itches, or the *development* into a different state of affairs that happens to be better for some reasons, but it need not be a move away from something worse that has come before, towards something better that comes ahead. In a similar fashion, my values are qualitatively different from those that I used to have when I was eighteen, but are they really so different and, crucially, so much better than before to be legitimately called an instance of *progress*? Isn't it just growth, or development? Isn't it all just a scratch? Wittgenstein seems to liken most of the work of philosophy to just that – a scratch: something momentarily satisfying, but not necessarily better than before (scratching can also make it worse).

"*Real* progress in philosophy, for Wittgenstein, is in another register entirely; it doesn't consist in what we are inclined to think of as progress. Real progress is discovering a cure: no longer needing to scratch the itch. Or simply no longer itching." (2016: 269).

Thus, real progress is not a developmental stage that is better than the one before, nor it is a growth: it is a *discovery*, and discoveries are much less frequent than we like to think.

“Our civilization is characterized by the word ‘progress’. Progress is its form rather than making progress one of its features”.

The illusion is that of thinking that societies and individuals *necessarily* move forward, that what we call ‘growth’ is also a ‘progress’, that is, that it necessarily leads to better outcomes than before. It may be so, but it is not a necessary condition. Overestimating change by calling it “progress” prevents us from realizing that it is very likely that what we are doing is actually scratching an itch, as opposed to curing the itch. Progress as a form is the illusion that we, as a society and as individuals, are moving forward; progress as a feature is thinking of progress as those less frequent discoveries that truly cure whatever it is that we are merely scratching. One may, then, wonder if there is any difference between “progress” and “revolution.” I will conclude this section by introducing current views on moral revolutions and argue that, in fact, real moral progress in a Wittgensteinian sense⁶⁸ may just be a moral revolution.

3.3 Moral Progress and Moral Revolutions

Real moral progress is, it seems, something substantially different from mere finger scratching. Wittgenstein’s talk of “discoveries” as real progress has a scientific flavor, and one that has been seriously considered by scholars who study moral revolutions. In a recent review, Klenk et al (2022) have examined the concept of moral revolutions as

⁶⁸ As well as a Kuhnian sense, as I will argue in Chapter II.

yet another form of radical moral change. Like moral progress, that of moral revolution is a notion that seems to belong to the sphere of the collective, or structural-level forms of change, rather than to the individual one,⁶⁹ with the notable exception of Walden (2015). Klenk et al (2022) notice that there are some interesting commonalities among scholars. First of all, for all of them, “what makes revolutions distinctive is the radicality, depth or fundamentality of the moral changes involved” (2022: 355). This aspect of radicality takes different shapes: for example, in Baker’s account moral revolutions involve a paradigm shift towards an “alternative paradigm incompatible with the established paradigm” (2019: 42). According to Eriksen (2019), the radical element is that moral revolutions are changes in fundamental aspects of morality. Lowe (2019) and Buchanan (2020) observe that moral revolutions involve changes in moral beliefs, such as gaining more conceptual depth (e.g. “justice”, Lowe 2019: 2; or starting to view non-human animals as having moral standing, Buchanan 2020). Concept deepening is also at play in Walden’s account of individual level moral revolutions, according to which a change is revolutionary when they are due to changes in “background framework that conditions our thinking about that subject” (Walden 2015: 283). All this is reminiscent of what some scholars identify as the necessary epistemic element of moral progress, that takes the form of concept deepening, gaining more moral knowledge, or better moral understanding.

Interestingly, Klenk et al (2022) notice that “[n]early all the accounts of moral revolutions surveyed invoke an analogy between Kuhnian scientific revolutions (Kuhn

⁶⁹ See Baker 2019, a revolution only occurs if a moral community embraces a paradigm shift; Eriksen 2019, according to whom moral revolution is change at the societal level in the form of shared practices; also, more implicitly, Kitcher 2012, Pleasants 2018.

1966) and moral revolutions” (357).⁷⁰ One striking element of the analogy is that of incommensurability: “moral aspects changed by the revolution are incomprehensible, unjustifiable or incompatible with the pre-revolutionary situation” (356),⁷¹ just like it is the case for scientific revolutions, whereby once the scientific paradigm has changed (e.g. the Copernican revolution) looking at the same phenomenon with the old paradigm strikes scientists as unthinkable. Notable disagreements, on the other hand, concern the speed at which revolutions take place⁷² and the nature of moral revolutions.⁷³ Moreover, it is not at all settled to what extent moral revolutions are to be distinguished from other kinds of moral change. For Pleasants (2018) and Baker (2019), for example, a moral revolution is a change in moral paradigm (Baker 2029: 21, 49), while any other kind of change is a change within a paradigm.

Finally, just as it was the case with moral progress, the relation between the individual and the structural level is yet to be settled. Kitcher (2012) doubts that individuals have causal efficacy when it comes to bringing about moral revolutions, although he recognizes the importance of “exceptional” and “brave” individuals in spreading awareness on the need for a revolutionary change.⁷⁴ On the other hand, Kumar and

⁷⁰ See Appiah 2010: 11, Kitcher 2012: 2, 333, Pleasants 2018, Baker 2019: 2–53, Lowe 2019. Questions as per whether we can talk of “normal morality” as Kuhn talked about “normal science”, thus as per whether the analogy is accurate, see Palmer & Schagrin 1978, Pleasants 2018

⁷¹ Baker notes that the earlier paradigm will come to be “unthinkable” after the revolution (2019: 42).

⁷² Appiah 2010 sees it as “a large change in a small time” (170), while for Pleasants 2018 it is slow, especially when it comes to individuals, Baker 2019 claims that generations may be needed for a revolutionary idea to be enacted by individuals, and Kitcher 2021 adds the element of arbitrariness.

⁷³ e.g. whether they are intentionally driven. See Baker 2019 and Pleasants 2018 on the role of “dissidents”. See Eriksen 2019, 2020 and Appiah 2010 for accounts that go in the direction of moral revolutions as not necessarily intentionally driven.

⁷⁴ See also Pleasants 2008 where he mentions “moral entrepreneurs”.

Campbell (2016), Walden (2015) and Lowe (2019) seem to allow for the impact that individual-level changes of some form may have on moral revolutions at the structural level. Interestingly, Klenk et al (2022) notice that “[w]hatever position the authors take on individuals’ role in revolutions, one point of agreement amongst the authors surveyed is that moral revolutions never happen by ‘cogent, rational arguments’ alone (Appiah 2010: 41, cf. Pleasants 2010)” (361).

Is there any conceptual difference between moral progress and moral revolutions? The obvious difference is that “progress” always implies a change in states of affairs that is *for the better* – otherwise, it would be moral “regress” – while revolutions, although mostly for the better, are not necessarily progressive, since they are not necessarily *for the better*.⁷⁵ Moreover, those of moral revolutions are, for the most part, descriptions of dynamics of change, while “moral progress” is often used to indicate also the content of changes.⁷⁶ We can, thus, conclude that that of moral revolutions is a possible dynamic of moral progress. In other words, when a moral revolution occurs and it is also for the better, then it is an instance of moral progress. Now, given that moral revolutions involve a fundamental change – e.g. a Gestalt switch, or a paradigm shift – and if we maintain the analogy with scientific revolutions, then moral revolutions do amount to moral discoveries. Thus, moral progress that comes about through a moral revolution is *real* moral progress in the Wittgensteinian sense: it is a cure, rather than the scratching of the finger.

We can see that moral progress in the form of moral revolution has the element of radicality that is also common to moral conversion; moreover, the fact that they come about as Gestalt switches suggests also that they are experienced as abrupt rather than as gradual, ladder-like changes. However, it may strike as odd to compare a phenomenon

⁷⁵ See Baker (2019) on Nazi’s medical revolution.

⁷⁶ With the risk of creating confusion, as noted by Sauer 2024.

that mostly takes place at the individual level – moral conversion, moral transformation – with one that generally indicates changes at the structural level, such as that of moral revolutions. In the next chapter I will argue that moral revolutions take place also at the agential level, and propose to call these particular forms of moral change “Moral Paradigm Shifts.” For now, let us conclude with a recap of what we have explored so far concerning how radical moral change is generally conceptualized, since it will be the starting point of next chapter’s thesis.

4. Conclusion

In this chapter I considered radical moral change in its three main conceptualizations: “moral conversion,” “moral transformation” and “moral progress.” Concerning moral conversion, I isolated the following elements: a moral conversion can be experienced as *abrupt* and *unplanned*, it triggers a radical reconfiguration of one’s moral outlook and, for this reason, it comes about as *disruptive* of what came before. It is moral, to the extent to which it involves a radical reconfiguration of values and ideals, independently of one’s preferred moral theory; it is a conversion, to the extent to which it is experienced as a point of turn – or, rather, of *no* return. There is, thus, an element of incommensurability, where the convert now has “new eyes” to see reality, and cannot go back to her former ones. Finally, it departs from religious conversion because what it is that ultimately triggers and guides the convert towards her new value system need not be a Deity: the phenomenology of religious conversion sees the change as exogenous, while that of moral conversion need not do so.

Concerning moral transformation, I analyzed two main conceptual threads: one that sees moral transformation as following an ascension path, the other that sees moral transformation as following a practical, activity-based path. The first one is a thread that sees its origins in Plato and that has been more recently revived by Wittgenstein and Murdoch: we saw that Wittgenstein showed – rather than explained – what it is to ascend as readers of the TLP, and finally contemplate the true form of reality. Murdoch, on her part, has taken seriously this contemplative element and made it the core of her moral philosophy, arguing that it is “loving attention” that allows one to climb the ascension ladder. Contemplation, therefore, is a moral activity which is inherently transformative. I then compared this thread with the activity-based one, which is that

advocated by Callard, and noticed that they go in two diametrically different directions: the ascension one requires selflessness, while the activity one requires self-centeredness.

Finally, concerning moral progress, I expanded on both its individual reading and its form as “moral revolution,” and concluded that while moral progress per se may just amount to growth, or development – which, according to Wittgenstein, is not *real* progress – moral revolution takes place as a Gestalt switch, or a change of paradigm, which is what makes it radical and abrupt, rather than gradual, or step-by-step.

There are some common patterns we can spot. To begin with, the radicality of the moral change can be experienced as either gradual or abrupt: epiphanic moral conversions are definitely abrupt, but what Dees called “conversion by evolution” is a narrative that, on the other hand, stresses just how gradual was the process that lead to his realizing one day that he had undergone a moral conversion. Similarly, moral transformations in the form of ascension do take the form of a gradual move to reach reality – or Beauty, or the Good – but it is also likely that when one succeeds in gaining closeness to reality, one is struck by an epiphany, and just like in a Gestalt-switch, the change is experienced as abrupt. Thus, epiphanic moral conversions may also be epiphanic moral transformations, and gradual moral conversions may also be gradual moral transformations. What distinguishes epiphanic from non epiphanic instances of conversion/transformation is the phenomenology of what it is that triggered the change. There is, however, another element to consider: changes can be experienced as exogenous or endogenous. The “epiphanic” element of some forms of moral change suggests that what it is that triggers the change is something external, something that is disclosed to the agent, that does not belong within the agent. This is mostly clear in moral conversions, since the “conversion” element is what suggests that there is something – as noted, not necessarily a Deity – that “calls” for our attention, but it is

also present in moral transformations as ascensions, since what it is that we are ascending to has a “magnetic force” that takes us outside of ourselves. We can thus conclude that when the change is experienced as abrupt and exogenous, moral conversion and moral transformation may be used interchangeably.

Finally, moral progress can also be understood as more or less gradual change for better states of affairs. In this case, when it comes to the individual level, moral progress may just be another name for gradual instances of moral conversion, or moral transformation. On the other hand, when moral progress takes the form of moral revolution, it comes about as a Gestalt-switch that shares its element of abruptness and disruptiveness with epiphanic moral conversions and epiphanic moral transformations. The obvious disanalogy is that moral revolution refers to structures rather than to what happens in agents.

In the next chapter I will propose a classification of moral change that revolves around the following threefold distinction: (i) whether the change is actively pursued or passively received, (ii) whether the change is gradual or abrupt, and (iii) whether the change is core or peripheral. Then, I will isolate a particular form of change, which is that of changes that are passive, abrupt and core (PAC), and analyze them through an analogy with Kuhn’s paradigm shifts in science. I will conclude that we they constitute what can be called “Moral Paradigm Shifts” and can be made sense of as agent-level instances of moral revolutions. In the very last section, I will provide a more informed analysis of moral conversion, moral transformation and moral progress, in light of my proposed categorization.

Plato (2015) *Simposio, Apologia di Socrate, Critone, Fedone*. Savino, E. (eds). Oscar Mondadori Classici.

Murdoch, I. (1992) *Metaphysics as a Guide to Morals* [MGM]. London, Penguin.

———, *The Sublime and the Beautiful Revisited* [SBR], pp. 261–286.

———, *The Idea of Perfection* [IP], pp. 299–336.

———, *On ‘God’ and ‘Good’* [OGG], pp. 337–365.

———, *The Sovereignty of Good over other Concepts* [SGC], pp. 363–385.

Wittgenstein

[CV] *Culture and Value*, trans. Peter Wince (University of Chicago Press, 1989).

[NB] *Notebooks 1914–1916*, ed. G. H. von Wright and G. E. M. Anscombe, trans. G. E. M. Anscombe, 2nd edn, (Oxford: Blackwell, 1979).

[PI] *Philosophical Investigations*, trans. G. E. M. Anscombe, 3rd edn (Oxford: Blackwell, 2001).

[TLP] *Tractatus Logico-Philosophicus*, trans. C. K. Ogden (London: Routledge and Kegan Paul, 1922).

[WSP] *Wittgenstein: Sources and Perspectives*, ed. C. G. Luckhardt (Ithaca, NY: Cornell University Press, 1979).

References

Appiah, A. (2010). *The Honor Code: How Moral Revolutions Happen*. New York, NY: Norton.

Archer, M.S. (1995). *Realist social theory: The morphogenetic approach*. Cambridge University Press.

Asher, K. (2017). *Literature, ethics, and the emotions*. Cambridge, England: Cambridge University Press.

Baker, R. (2019). *The Structure of Moral Revolutions: Studies of Changes in the Morality of Abortion, Death, and the Bioethics Revolution*. Cambridge, MA: MIT Press.

Bien, T. (2004). Quantum change and psychotherapy. *Journal of Clinical Psychology*, 35 (4), 493–501.

Blum, L., "Iris Murdoch", *The Stanford Encyclopedia of Philosophy* (Fall 2023 Edition), Edward N. Zalta & Uri Nodelman (eds.), URL = <https://plato.stanford.edu/archives/fall2023/entries/murdoch/>.

Buchanan, A.; Powell, R. (2018). *The Evolution of Moral Progress: A Biocultural Theory*, Oxford University Press.

Buchanan, A. (2020). *Our Moral Fate: Evolution and the Escape from Tribalism*. Cambridge, MA: MIT Press.

Callard, A. (2018). *Aspiration: The Agency of Becoming*. New York: Oup Usa.

Callard, A. (2020). 'Transformative Activities', in Enoch Lambert, and John Schwenkler (eds), *Becoming Someone New: Essays on Transformative Experience, Choice, and Change* (Oxford, 2020; online edn, Oxford Academic, 20 Aug. 2020), <https://doi.org/10.1093/oso/9780198823735.003.0009>

Caprioglio Panizza, S. (2022). *The Ethics of Attention. Engaging the Real with Iris Murdoch and Simone Weil*. New York: Routledge.

Caprioglio Panizza, S., Vaccarezza, M. S. (2024). Moral Transformation as Shifting (Im)Possibilities. *The Journal of Ethics*:1-16.

Chang, R. (2015). Transformative Choices. *Res Philosophica* 92 (2):237-282.

Chappell, S. G. (2022). *Epiphanies: An Ethics of Experience*. Oxford: Oxford University Press.

Cordner, C. (2016) Lessons of Murdochian Attention. *Sophia*. 55: 197–213.

Cottingham, J. (2012). Human nature and the transcendent. *Royal Institute of Philosophy Supplements* , 70 , 233–254.

Dees, R. H. (1996). Moral conversions. *Philosophy and Phenomenological Research* , 56 (3), 531–550.

Destrée, P.; Giannopoulou, Z. (2017) Eds. *Plato's Symposium. A Critical Guide*. Cambridge: Cambridge University Press.

Dickens, C. (1995). *A Christmas carol and other stories*. New York: Modern Library.

Eriksen, C. (2019). The dynamics of moral revolutions: prelude to future investigations and interventions. *Ethical Theory and Moral Practice* 22: 779–92.

Gandhi, M. K. (1927). *An Autobiography; or the Story of My Experiments with Truth*. Translated by Mahadev Desai. Ahmedabad, India: Navajivan Publishing House.

Gibbs, J. C. (2014). *Moral development and reality: Beyond the theories of Kohlberg, Hoffman, and Haidt* (3rd ed.). Oxford, England: Oxford University Press.

Griffin, E. (1980). *Turning: Reflections on the experience of conversion*. Garden City NY: Doubleday.

Hämäläinen, N. (2019). “Which Void? (MGM chapter 18)”, in Dooley and Hämäläinen 2019: 261–275.

Hermann, J. (2019). The dynamics of moral progress. *Ratio*, 32(4), 300-311.

Hugo, V. (1983). *Les Misérables*. Einaudi.

Irvine, W. B. (2015). *Aha! The moments of insight that shape our world*. Oxford, England: Oxford University Press.

James, W. (1958). *The varieties of religious experience: A study in human nature*. New York, NY: Mentor Books.

Jamieson, D. (2002). Is there Progress in Morality? *Utilitas* 14(3):318–338.
<https://doi.org/10.1017/S0953820800003630>.

Jonas, M. (2015a). Education for epiphany. The case of Plato's *Lysis*. *Educational Theory*, 65 (1), 39–51.

Jonas, M. (2015b). Habituation and role- modeling in Plato's thought. Paper presented at a Jubilee Centre for Character and Virtues Conference, Oriel College, Oxford, January 8–10. Retrieved August 11, 2016, from www.jubilee-centre.ac.uk/userfiles/jubileecentre/pdf/conference-papers/Varieties_of_Virtue_Ethics/Jonas_Mark.pdf

Kitcher, P. (2012). *The Ethical Project*. Cambridge, MA: Harvard University Press.

Klenk, M., Sauer, H. (2021). "Moral Judgement and Moral Progress: The Problem of Cognitive Control", *Philosophical Psychology*, 34(7), 938–961.

Klenk, M.; O'Neill, E.; Arora, C.; Blunden, C.; Eriksen, C.; Frank, L. & Hopster, J. (2022). Recent Work on Moral Revolutions. *Analysis* 82 (2):354-366.

Kohlberg, L., & Ryncarz, R. A. (1990). Beyond justice reasoning: Moral development and consideration of a seventh stage. In C. N. Alexander & E. J. Langer

Kraut, R. (1973). "Egoism, Love and Political Office in Plato", *Philosophical Review* 82: 330–44.

Kraut, R. (1999). Ed. *The Cambridge Companion to Plato*. Cambridge: Cambridge University Press.

Kraut R. (2017). Eudaimonism and Platonic erōs. In: Destrée P, Giannopoulou Z, eds. *Plato's Symposium: A Critical Guide*. Cambridge Critical Guides. Cambridge University Press :235-252.

Kristjánsson, K. (2018). "Epiphanic Moral Conversions", in *Self-Transcendence and Virtue*, Jennifer A. Frey and Candace Vogler (eds), Routledge, pp. 15-38.

Kumar, V., Campbell, R. (2022). *A Better Ape: The Evolution of the Moral Mind and How it Made Us Human*, Oxford University Press.

Leuba, J. H. (1896). "A Study in the Psychology of Religious Phenomena." *The American Journal of Psychology* 7, no. 3: 309-385.

Lewis, C. S. (1955). *Surprised by Joy: The Shape of My Early Life*. New York: Harcourt Brace Jovanovich.

Liberti, M. (2024). Neoptolemus and Huck Finn Reconsidered. Alleged Inverse akrasia and the Case for Moral Incapacity. *J Value Inquiry*. <https://doi.org/10.1007/s10790-024-09981-w>

Lowe, D. (2019). The study of moral revolutions as naturalized moral epistemology. *Feminist Philosophy Quarterly* 5: 1–15.

Macklin, R. (1977). Moral progress. *Ethics*, 87(4), 370-382.

MacLaughlin, A. J. (2008). Narratives of hope: A philosophical study of moral conversions . Ph.D. diss. Loyola University, Chicago.

Madva, A. (2016). A plea for anti-anti-individualism: How oversimple psychology misleads social policy. *Ergo*, 3.

Malcolm, N., Wright, G. H. von (1971). Ludwig Wittgenstein A Memoir With A Biographical Sketch. Oxford University Press.

Maslow, A. (1964). Religions, values, and peak- experiences . Harmondsworth, England: Penguin Books.

Moody-Adams, M.M. (1999). The idea of moral progress. *Metaphilosophy*, 30(3), 168-185.

Moore, George Edward (1903). *Principia ethica*. Mineola, N.Y.: Dover Publications. Edited by Thomas Baldwin.

Morris, M. (2008). Routledge Philosophy Guidebook to Wittgenstein and the *Tractatus*. Routledge.

Nguyen, C. T. (2019). Games and the art of agency. *Philosophical Review* 128 (4):423-462.

Obdrzalek, S. (2010). "Moral Transformation and the Love of Beauty in Plato's Symposium", *Journal of the History of Philosophy*, 2010 (Vol. 48), 415-44.

Paul, L. A. (2014). *Transformative experiences*. Oxford, England: Oxford University Press.

Pleasants, N. (2018). The structure of moral revolutions. *Social Theory and Practice* 44: 567–92.

Rambo, R. L., & Farhadian, C. E. (Eds.). (2014). *The oxford handbook of religious conversion*. Oxford: Oxford University Press.

Read, R. (2016). Wittgenstein and the Illusion of 'Progress': On Real Politics and Real Philosophy in a World of Technocracy. *Royal Institute of Philosophy Supplement* 78:265-284.

Robjant, D. (2012). "The Earthy Realism of Plato's Metaphysics, or: What Shall We Do with Iris Murdoch?", *Philosophical Investigations*, 35(1): 43–67. doi:10.1111/j.1467-9205.2011.01455.x

Rorty R. (1999). Human rights, rationality and sentimentality. In O. Savic (ed.), *The Politics of Human Rights*, Verso, 67-83.

Sartre, J-P. (1964). *Nausea*. Translated by Lloyd Alexander. New York: MJF Books.

Sauer, H. (2023), *Moral Teleology: A Theory of Progress*. Routledge.

Schinkel, A. (2016). *Morally formative experiences* (manuscript under construction).

Schinkel, A., de Ruyter D.J. (2017). Individual moral development and moral progress, *Ethical Theory and Moral Practice*, 20(1), 121–136.

Schnider, A. (2008). *The Confabulating Mind: How the Brain Creates Reality*. Oxford University Press.

Severini, E. (2021). Moral Progress and Evolution: Knowledge Versus Understanding. *Ethical Theory and Moral Practice*, 24(1), 87-105.

Singer, P. (2011). *The expanding circle: Ethics, evolution, and moral progress*. Princeton University Press.

Slote, M. (2007). *The Ethics of Care and Empathy*. Routledge.

Songhorian, S., Guma, F., Bina, F., Reichlin, M. (2022). Moral Progress: Just a Matter of Behavior?, *Teoria*, 42(2), 175-187.

Spinney, O. T. (2023). "Religious conversion, philosophy, and social science", *International Journal for Philosophy of Religion*, 94:139–149, <https://doi.org/10.1007/s11153-023-09873-2>

Taylor, C. (1989). *Sources of the self: The making of the modern identity*.

Vaccarezza, M. S. (2018). *Virtù esemplari. L'etica Tommasiana tra Neoplatonismo e Aristotelismo*. *Etica & Politica / Ethics & Politics* , XX, 2018, 2, pp. 15-30.

Vaccarezza, M. S. & Niccoli, A. (2022). Let the donkeys be donkeys: in defense of inspiring envy. In Sara Protasi (ed.), *The Moral Psychology of Envy*. Rowman & Littlefield Publishers. pp. 111-127.

Walden, K. (2015). Art and moral revolution. *The Journal of Aesthetics and Art Criticism* 73: 283–95.

White, F. (2014). *Becoming Iris Murdoch*, Kingston-upon-Thames: Kingston University.

Woodruff, D. (1999). *Tractarian Mysticism. Moral Transformation through aesthetic contemplation in Wittgenstein's early philosophy*. PhD diss. The University of Illinois.

Yu, J. (2007). *The ethics of Confucius and Aristotle: Mirrors of virtue* . London, England: Routledge.

Chapter II. Moral Paradigm Shifts

This chapter has the following purposes:⁷⁷

1. Propose three parameters around which to categorize different forms of radical moral change;
2. Isolate one particularly meaningful form of radical moral change: moral paradigm shifts.

I will then advance the following argument:

(i) just like scientists make sense of observed phenomena through scientific paradigms, agents make sense of moral phenomena through moral paradigms

(ii) scientific revolutions occur as scientific paradigm shifts

(iii) scientific paradigm shifts are PAC changes

=> PAC changes are paradigm shifts

=> moral PAC changes are Moral Paradigm Shifts (MPS).

I will then stress that MPS are important because (a) they are the form of real moral progress, and (b) because they place demands on moral theories, as we will see in Chapter III.

Chapter I kicked off on the following observation: people do change in radical and morally significant ways. Maybe something happens; they survive an accident, they fall in love, someone dies, someone is born, their life does not go as planned, they stumble upon books, stories that trigger something; or maybe nothing happens, they just wander

⁷⁷ This chapter was drafted under the generous supervision of prof. Uriah Kriegel (Rice University).

and suddenly they are struck by a realization. Maybe they hear something: threads of dialogues, or a song, or the wind. All this to say that just how any of these possibilities can in fact lead to significant moral change is hard to make sense of.

As a first attempt to shed some light on the phenomenon of radical moral change, I offered a conceptual map that unfolds around three main conceptualizations of radical moral change: moral conversion, moral transformation and moral progress. I concluded by trying to draw a preliminary sketch of how these forms of change intertwine, if anything, with one another, but the lack of conceptual tools to make sense of moral change has made it pretty tentative.

In this chapter I propose a way to classify forms of moral change (section 1), depending on the following three elements: (i) whether the change is actively pursued or passively received, (ii) whether the change is gradual or abrupt, and (iii) whether the change is core or peripheral. I then narrow the scope of this dissertation to analyzing one particular form of change, which is that of changes that are passively received, abrupt and core (PAC), which, after drawing an analogy with Kuhn (section 2), I call Moral Paradigm Shifts (MPS). In the final part I will apply the classification of moral change to the three conceptualizations considered in Chapter I in order to look for MPS in the literature considered. In the next and final chapter I will explore MPS' importance for moral theories and attempt to make room for the possibility of MPS in the moral theory that is most inhospitable for radical moral change; namely, Virtue Ethics.

1. Conceptual Framework

In this part of the dissertation I will begin by suggesting a conceptual framework to make sense of the dynamics of moral change, which distinguishes different forms of moral change depending on the degree to which they are actively pursued or passively

received (1.1), experienced as gradual or abrupt (1.2) and whether they take place at a core or at a superficial level (1.3). I will then isolate one form of change that I believe to be particularly significant, which is that of PAC changes; that is, of changes that are passively received, experienced as abrupt and that take place at the level of core commitments. In the second section I will provide an analogy between PAC changes and Kuhn's account of paradigm shifts in science, and conclude that we can legitimately call PAC changes as "Moral Paradigm Shifts". I will conclude by considering those instances of moral conversion, moral transformation and moral progress that are moral paradigm shifts.

1.1 Active and Passive

A change can occur in virtue of a decision, a wish or a plan – that is, change can occur in virtue of subjects actively pursuing it in some fashion, or in virtue of subjects deciding to be open to the possibility of change. In this case, the subject of change is active.⁷⁸ On the other hand, change can occur in an unplanned, non-pursued, unexpected way – that is, change can occur while the subject is merely passive to it. In this case, the subject of change is passive.⁷⁹ In the former case, the changed subject can claim some sort of ownership over the process of change, while in the latter they cannot, since the change is acted upon the agent, not by the agent: it just happens *to* them, not in virtue of them.⁸⁰

⁷⁸ Or, an agent of change.

⁷⁹ See Taylor 1982 for skepticism concerning the meaningfulness of this distinction.

⁸⁰ This distinction can also be framed in terms of agency/patency. See Callard (2020), where she opts for an agency/patency distinction that revolves around the process of learning, according to which agent learning is a

Active forms of change are, to some extent, planned. Deciding to take a course in philosophy to gain deeper understanding is an active form of change, since the subject willingly puts herself in a potentially transformative situation, or decides to challenge her stances on some matters. In a similar fashion, deciding to hang out with a wide variety of people, or to live abroad for some time are good indicators that one is, at least in principle, open to the possibility of changing; thus, if they indeed undergo a process of change, they do so actively. In all these situations, the subject cannot predict whether she will actually change or not, nor how and to what extent she will indeed change; but change in some direction is a possibility, and one they can will and act upon. On the other hand, forms of change that are passive are not planned, nor were they present as a possibility to the subject before the change had taken place. The best way to make sense of this distinction is, thus, in terms of practical possibilities: in active cases, changing in some direction is a possibility, while in passive cases it is not, until it happens.

1.2 Abrupt and Gradual

An abrupt change is one that is experienced as a sudden disruption of the some previous state of affairs.⁸¹ A gradual change, on the other hand, is one that is experienced as coming about through a series of steps, in which every step naturally arises from the one that came before, without any sort of perceived “break”.⁸²

process in which the agent takes active part by inquiring, while patient learning process is merely being passively informed of something.

⁸¹ In Chapter I we saw Kristjánsson 2018 on the abruptness of what he calls “epiphanic moral conversions”; MacLaughlin on “moral conversions” 2008 and Bien 2004 on “quantum psychological changes”.

⁸² Something like Kohlberg’s stages (1981), or neo-Aristotelian moral development (Sanderse 2015). See also Chapter III of this dissertation.

Now, it is important to note that this distinction is phenomenological, since it has to do with how people *experience* change, and not with whether change has *objectively* happened abruptly or gradually. An external observer could examine another person's abrupt change and always find some possible gradual explanations in her life story; nonetheless, the subject of change could still experience abruptness, even if somewhere in her life there had been a moment that could have "prepared" the change and, thus, may suggest continuity. To picture this, imagine Nora, a girl who grew up in a particularly racist community and who, one day, finds herself incapable of bullying her black classmates anymore. What happened is that she has undergone some sort of morally relevant change which she may have experience as more or less abruptly. We can imagine her gradually distancing herself from her family's racist views, beginning to question the moral postulates that were given to her (e.g. "non-white people are uncivilized and dangerous"), to the point that, one day, she just does not find her former racist views as action-guiding anymore. If we asked her about her change, she may reply that she could easily pinpoint the steps that brought her to that day in which she stopped endorsing racist views: "I was puzzled by my mother's racist remark, so I decided to ask her what she meant and I realized that she clearly did not have a logical answer; there was that other time we needed an extra player for a football game and I asked Alagie, my classmate from Gambia, to join us out of necessity. It was fun. After that I decided to talk to him and his friends, invite them more often ...". In this case, Nora has experienced her change as a series of gradual steps.

However, we can also imagine Nora providing the following testimony: "I always found my mother's racist remarks funny; not once they made me feel uneasy, and I even refused to play football that day my friends invited Alagie. Yet there I was, the day of Kya's philosophy oral test, when the teacher was humiliating her because she could not remember the phases of consciousness in Hegel's phenomenology. I am not sure why,

but I took the philosophy book, looked for the page on Hegel’s phenomenology and whispered to Kya’s ears. Upon feeling the urge to help her, I realized something had changed inside me”. In this case Nora has experienced her change as abrupt: nothing that came before made her suspect that she was about to change, it came as utterly sudden and unexpected. As mentioned before, it is likely that an external observer, or Nora herself some time afterwards, may look at her story and manage to pinpoint some key moments that anticipated the change, making it look more gradual than she may have thought – but the fact that she was not aware of these key moments made it so that she did not experience these key moments as proper “steps”. Hence, the felt abruptness of her change.

1.3 Peripheral and Core

Finally, if we consider agent’s morally significant actions and beliefs as being in some sort of relation with what they deem important (e.g. values, principles, ideals), then we can draw a third distinction: moral change can impact what is of peripheral importance to the agent, or what is of core importance to the agent.⁸³ Something that is peripherally important does not impact the agent’s overall structure of action-guiding values and principles, but only some non fundamental aspects of it. So-called “changes of opinion” or “changes of mind” could classify as peripheral changes, since they do not need to imply a structural re-arrangement of action-guiding values and principles. On the other hand, a moral change that qualifies as core is one that impacts those values that are

⁸³ See Panizza and Vaccarezza (2024).

fundamental to the agent's moral system, so much so that that when these values change, a massive readjustment of the value system is in need.⁸⁴

Imagine Marc, a possessive, violent person who, upon talking to his sister Tina, decides to change his attitudes towards his wife in order to be "a better husband". Tina knows that Marc has violent dispositions, so she tries to hint at the fact that he should at least consider treating his wife with respect. Marc ponders on his sister's words and starts washing the dishes instead of leaving them around the house. However, his values of the violent, proud macho remain untouched; that is, he merely changed one particular instance of behavior, not his fundamental commitments. He will still be unquestionably prone to violence and to degrading his wife in other manners.⁸⁵

One way to make sense of this peripheral/core distinction is by framing it in terms of moral self-identity (Blasi 1984). There are certain values and commitments that are core to who we are as moral agents, while some that are peripheral; that is, they do not influence the whole structure of who we are as moral agents. Someone whose core value is friendship will shape her entire value system around friendship, and where she to be in the position to act contrary to such value (say, for the greater good), she would feel as if she was betraying herself.⁸⁶ On the other hand, going against values that are

⁸⁴ This seems to be what Kristjánsson has in mind when talking about changes that are "spontaneous and dramatic, involving radical reconfigurations of mental structures" (2018: 18).

⁸⁵ Here I am not suggesting that one engages in violent behavior *because* one values violence. One may engage in violent behavior for a number of complex psychological reasons that have nothing to do with what the agent values and what she does not value (e.g. what she has suffered in the past). I am talking about a specific case in which the agent both engages in violent behavior *and* does not see anything in principle wrong with it. A moral change that happens at the central level would imply questioning and possibly revising the second part.

⁸⁶ See Lapsley 2008 on a discussion of the threat of self-betrayal in Blasi's account of moral self-identity.

merely peripheral is much less dramatic. One may have honesty as a peripheral value, thus finding it action-guiding just at times, but not always; failing to pursue honesty, in this case, would not threaten one's integrity as a moral agent. In our previous example, Marc's change involved a minor change in behavior, so we classified it as peripheral; if Marc's change involved his commitment to violence and machismo, then it would have been core and, thus, likely to involve not only him washing the dishes, but also revising other inherently violent habits that he may possess.⁸⁷

Now, one may wonder whether peripheral changes do classify as *moral* changes at all: are they not morally trivial by definition? Before proceeding, we need at least some to discern what it is that makes moral changes *moral* in the first place. As we saw in Chapter I, Alfredo MacLaughlin (2008) distinguishes between two complementary senses of "morality", namely, morality as considerations of Right or Wrong (RW) and morality as considerations on the Meaning of Life (ML), in virtue of which it is possible to draw distinctions between moral and non-moral changes:

(RW) if we understand morality as having to do with matters of right and wrong, then morally relevant changes are those changes in virtue of which the agent changes "*with regard to her/his existential involvement in the task of acting rightly or wrongly, of doing what ought to be done*" (2008: 83). Crucially, he adds that "[these] change[s] may involve both big, consequential decisions and the small, perhaps half-thought actions of daily life" (ibid).

(ML) if we understand morality as having to do with the meaning of life, then morally relevant changes are those changes in virtue of which the agent changes with regard to her/his existential involvement in the task of living a meaningful life, of engaging in flourishing activities. As it is the case for (RW), these changes may involve both big,

⁸⁷ As mentioned in Chapter I, MacLaughlin defines core changes as "structurally significant" (2008: 127).

fundamental decisions concerning the meaning of life and the small, perhaps half-thought actions of daily life.

Thus, any change that regards agents' concerns with doing the right thing, what ought to be done, and/or what they think it is that should guide their actions towards a flourishing life, is a morally relevant change. These changes can have degrees of moral significance, depending on whether they come about at the level of "small, perhaps half-thought" everyday actions, or at the more fundamental, "structural" level of core commitments to what is right/wrong, what ought to be done and what counts as a life worth living. In the former case, the change is *peripheral*; in the latter case, the change is *core*. In the former case, the change is only very mildly morally significant; in the latter case, the change is highly morally significant. In both cases, the change is morally relevant.

1.4 Triplets

Moral change can come about as more or less active or passive, more or less gradual or abrupt, and more or less peripheral or core. These three possibilities combine together to create proper dynamics of change:

Active, Abrupt, Core (AAC)

The change is actively pursued, or the agent is open to the possibility that change may come about; the change is experienced as abrupt and it takes place at the level of what is core to the agent. Think of therapy groups, where people actively⁸⁸ place themselves in a position to discuss their experiences and question what it is that has been guiding them until that moment. Take possessive Marc once again: he may wish to stop showing possessiveness and jealousy towards his partner and, therefore, decide, or surrender to

⁸⁸ not always, of course.

encouragement, to participate in a therapy group. Upon hearing another participant's story, he suddenly switches his perspective on his partner: like in a Gestalt switch, he stops seeing her as his property and switches to seeing her as an independent human being. Something deep has changed, to which he now has to readjust. He comes back home and does not really know what to do anymore, he has no guidance anymore, he only knows he wishes his partner to be happy.

Active, Abrupt, Peripheral (AAP)

Change that is willingly pursued, or where the agent is open to the possibility that the change may come about, that is experienced as sudden and that takes place at the level of what is merely peripheral to one's moral structure. Imagine Mark who takes on the suggestion to go to group therapy. Upon hearing another participant's story, Marc finds himself questioning whether he should, after all, leave all the housework to his partner; from that day on, he starts helping his partner with the dishes and the laundry, but never once does he question whether his possessiveness and jealousy are dangerous elements of the relationship. This is a change at the level of morally insignificant, although relevant, features of Marc's moral identity.

Active, Gradual, Core (AGC)

Change that is willingly pursued, or where the agent is open to the possibility that change may come about; that is experienced as the result of gradual steps, and that takes place at the level of what is core to the agent and that, thus, calls for massive readjustment. Take Marc again: we imagined his Gestalt switch as abrupt in AAC, while here we can imagine him experiencing the change as a slow, step-by-step process, where session after session he comes to question his previous possessive outlook towards his partner, until one day he realizes that some deep change has taken place

inside of him. This change calls for massive readjustment of values, priorities and, in general, whatever it is that guides his actions.

Active, Gradual, Peripheral (AGP)

Change that is willingly pursued, or where the agent is open to the possibility that change may come about, that is experienced as the result of gradual steps and that takes place at the level of what is peripheral to the agent's moral structure and, thus, does not call for massive readjustment. Our Marc from AAP may well have experienced the realization that he should wash the dishes as gradual rather than abrupt; that is, as the result of a number of testimonies or group exercises that he may have heard or participated in.

Passive, Abrupt, Core (PAC)

Change that is not willingly pursued, or where the agent is not open to the possibility of change, that is experienced as sudden, unexpected, and that takes place at the level of what is core to the agent's moral structure. Imagine Nora again: she does not willingly pursue the possibility of endorsing non-racist values, but one day she happens to read, say, *The Adventures of Huckleberry Finn* and suddenly undergoes a Gestalt switch, where she cannot look at her non-white classmates according to her family's racist values anymore. This calls for massive readjustment of what it is that guides her actions.

Passive, Abrupt, Peripheral (PAP)

Change that is not willingly pursued, or where the agent is not open to the possibility of change, that is experienced as gradual, step-by-step, and that takes place at the level of what is peripheral to the agent's moral structure. Nora reads *Huckleberry Finn* and, suddenly, she realizes that she should spend less time scrolling social media and more

time reading books. This change does not call for massive readjustment of her core values, thus, it is not significant in this sense, while still morally relevant, since it has to do with what it is that guides her actions, or with her conception of a good life.

Passive, Gradual, Core (PGC)

Change that is not willingly pursued, or where the agent is not open to the possibility of change, that is experienced as gradual, step-by-step, and that takes place at the level of what is core to the agent's moral structure. Nora's Gestalt switch concerning how she sees and treats her non-white classmates can be experienced also as a slow, step-by-step process, where, for example, she may start from realizing how one particular scene of *Huckleberry Finn* can be relevant to her situation as well⁸⁹ and, thus, how that may apply to her too and not just to the characters of the book. She may then work on her massive revision of values in a slow, continuous way.

Passive, Gradual, Peripheral (PGP)

Change that is not willingly pursued, or where the agent is not open to the possibility of change, that is experienced as gradual, step-by-step, and that takes place at the level of what is merely peripheral to the agent's moral structure. Nora finds herself, gradually and continuously, reaching for more books rather than for time spent scrolling social media.

⁸⁹ Maybe the famous scene where Huck Finn does not give Jim away to the slave hunters. It is all the more interesting to note that a racist person could empathize with Huck Finn both when he thinks he should give Jim away, but also when he does not do so in the end, because Jim is his friend.

In this dissertation I am interested in one particular triplet, which is that of changes that are Passive, Abrupt and Core (PAP); that is, changes that are not willingly pursued, or where the agent is not open to the possibility of change, that are experienced as sudden disruptions of what came before and that take place at the level of what is core to the agent's moral structure. The uniqueness of this form of change can be appreciated in contrast with other forms of change listed above.

Firstly, while AAC (Active, Abrupt and Core) changes come about as somewhat unsurprising, given that they are actively pursued by the agent, or are those changes towards which the agent would in principle be open about, PAC (Passive, Abrupt and Core) come across as striking, unexpected. Our AAC example was Marc going to group sessions to be a good partner: he may not be able to predict the extent to which he will succeed in becoming a good partner, nor how resistant he may be to revise some of his attitudes, but he is at least in principle open to the idea of turning into a better partner, with all that it make take. Our PAC example, on the other hand, is Nora, who has no plan whatsoever to question her racist values, nor does she have any doubt concerning their action-guidance. She does not read *Huckleberry Finn* in order to be challenged on her core beliefs; yet, she ends up re-arranging her entire value system. Thus, PAC are more similar to epiphanies rather than to actively pursued changes.⁹⁰

Secondly, we can see that PGC (Passive, Gradual and Core) are changes where the only element of difference with PAC is that the change is experienced as gradual rather than as abrupt. Our PGC example was Nora simply taking steps in order to operate the re-framing of her core values. The phenomenological difference is that in PAC the re-framing of values is experienced as a Gestalt switch, while in PGC it takes the form of a gradual development from one frame to another.

⁹⁰ See Chappell 2022.

Finally, PAP (Passive, Abrupt and Peripheral) changes differ from PAC only in terms of where the change takes place, whether at the core or at the peripheral level. In PAP, Nora undergoes an unplanned, unexpected change towards reading more books as opposed to scrolling social media: to the extent to which reading more books rather than scrolling social media may be understood as a duty to enrich one's cultural background, or as, generally, a way to live a better life, this type of change is morally relevant, yet not that significant. In short, she is not radically revising her core values and principles (although, if she starts reading more often, she may eventually do so); she is just changing her habits, without questioning what guides her morally significant actions. On the other hand, in PAC, the very same thing shakes Nora to the core, to the point that she finds herself incapable of bullying her classmate with the same nonchalance as she used to have. Thus, we can preliminary conclude that PAC changes are epiphanic forms of change that are experienced as disruptive of previous internal moral status quo and that call for massive re-adjustment of one's value system.

These differences in elements may strike as nuances; in fact, as mentioned before, distinctions should not be taken too sharply. A change may be unconsciously planned but consciously unplanned; it may be experienced as abrupt, but develop gradually; it may take place at the level of core values, yet not change much in terms of attitudes and habits. All these and possibly more factors make categorical distinctions hard to draw. On the other hand, whether the change is planned or not, whether it is experienced as abrupt or gradual and whether it shakes one's core completely or not are possibilities that do allow for different forms of change; consequently, the categorization proposed helps us making sense of dynamics of change such as those of Moral Conversion, Moral Transformation and Moral Progress, which I will consider once again in section 3. In what follows, I will expand on the dynamics of PAC changes and advance an analogy with Kuhn's account of paradigm shifts.

2. Moral Paradigm Shifts

We have seen that PAC forms of change are unique in their being a sort of epiphanic Gestalt switch, where the agent suddenly experiences a radical re-framing of what is in front of her; that is, where the agent radically shifts her moral outlook. This form of change is particularly significant for a moral theory to make sense of, since it questions the very idea of improvement or progress as being a process that is linear and gradual; that is, it goes in the direction of Wittgenstein's critique that we saw at the end of Chapter I, according to which real progress is revolutionary, while fake progress is more akin to improved state of affairs that could just amount to fingers not itching anymore. However, we can further appreciate the significance of PAC changes by looking at the way they resemble what Kuhn has famously called "paradigm shifts" in science. In what follows I will consider the revolutionary dynamics of scientific paradigm shifts and argue that, in virtue of its resemblance with PAC changes, we can legitimately call PAC changes "moral paradigm shifts".

Now, claiming that there are such things as moral paradigm shifts implies that there is such a thing as a moral paradigm in the first place. It is important to note that "paradigm shift" is a term that does not commonly refer to processes that happen to agents, but rather to structures. In what follows, we will see that we have good reasons to think that paradigm shifts take place also at the agential level.

I will start by stressing that if scientists want to make sense of observed phenomena, they need a conceptual framework that defines the rules of meaningful engagement with such phenomena. Similarly, if agents want to make sense of moral situations, they need

a conceptual framework that define the rules of meaningful engagement with such situations. I will consider Kuhn's discussion on the necessity of paradigms for scientific practice and conclude that, in a similar fashion, paradigms are necessary for moral practice. Thus, there are such things as moral paradigms.

Secondly, I will consider Kuhn's account of scientific paradigm shifts and analyze it in light of the proposed categorization of forms of change: active/passive, abrupt/gradual, peripheral/core. If scientific paradigm shifts share the same parameters of the kinds of moral changes that I consider here – that is, if they are passive, abrupt and core – then we can legitimately call those kinds of moral changes “moral paradigm shifts”. I argue that, in fact, they do.

2.1 Scientific Paradigms and Moral Paradigms

Kuhn defines “normal science” as “research firmly based upon one or more past scientific achievements, achievements that some particular scientific community acknowledges for a time as supplying the foundation for its further practice” (1970: 183). These achievements, he continues, present two characteristics: they are (i) “sufficiently unprecedented to attract an enduring group of adherents away from competing models of scientific activity”, and (ii) “sufficiently open-ended to leave all sorts of problems for the redefined group of practitioners to resolve” (ibid.). These achievements go to construe what Kuhn calls “paradigms”; that is, “models from which spring particular coherent traditions of scientific research” (ibid.). That is to say, that whenever scientists' research is based on shared paradigms, scientists are “committed to

the same rules and standards for scientific practice” (ibid.).⁹¹ Thus, broadly speaking, paradigms are “background frameworks” (Walden 2015) that condition the way we interact with a subject matter. They provide the conceptual scheme through which we make the subject matter intelligible, and provide the rules through which we can interact with the subject matter. Crucially, they are employed as models to solve puzzles, so much so they become tradition.⁹² In particular, a puzzle is solved by testing the scientist’s best explanations against those theoretical postulates that are provided by the paradigm employed:

“[i]n no usual sense, however, are such tests directed to current theory. On the contrary, when engaged with a normal research problem, the scientist must premise current theory as the rules of his game. His object is to solve a puzzle, preferably one at which others have failed, and his current theory is required to define that puzzle.... Of course the practitioner of such an enterprise must often test the conjectural puzzle solution that his ingenuity suggests. But only his personal conjecture is tested.” (London Lecture pp. 4–5).

This suggests that in order to solve puzzles, scientist need to take the rules suggested by the current paradigm as postulates, and then test their intuitions in light of such

⁹¹ He also explains paradigms through an analogy with language learning: “[pupils] learn for example, to recite, amo, amas, amat, amamus, amatk, amant, and they then use that standard form to produce the present active tense of other first conjugation Latin verbs. The usual English word for the standard example employed in language teaching is ‘paradigms’, and my extension of that term to standard scientific problems like the inclined plane and conical pendulum did it no apparent violence.” (Kuhn ‘Preface’, in *The Essential Tension*, p. xix 1977, in Cedarbaum 1983).

⁹² It appears that “[f]ew books have been as widely read and as widely misunderstood as *Structure*.” (Paradigms, Cedarbaum 1983).

postulates. Thus, paradigms are the prerequisites of normal science, since without a paradigm, there would be no postulates in light of which scientists could test their intuitions.

If we borrow Kuhn's account of "paradigms" and transpose it to the moral domain, we get the following picture: moral paradigms are those conceptual schema through which we make sense of morally charged situation. The community in question is not that of scientists, but that of moral agents, and the puzzles are morally relevant dilemmas of everyday life. We approach them through already established rules, we try to make sense of them in a way that can be understood by other members of our community of moral agents, and try to address them in an intelligible, rational way. In literature, moral paradigms are what generate our "moral outlook". Now, there are a number of ways to make sense of "moral outlooks". One story is offered by McDowell, who defines it as "a specific determination of one's practical rationality" (1998: 58) which is a definition that, as noted by Vigani, has "both evaluative and affective aspects" (2018: 2). According to McDowell, what distinguishes the virtuous from the *enkrates* is precisely her moral outlook: that of the virtuous is one that sees only virtuous reasons for actions in each situation, while that of the *enkrates* is one that sees both virtuous and non-virtuous reasons for actions in each situation, causing the typical *enkrates* distress of seeing what virtue requires, yet desiring something different.⁹³ In other words, the moral outlook in this sense is what allows the virtuous agent to see only the virtuous requirements of the situation, and the *enkrates* to see a number of possibilities for action.

⁹³ See Vigani 2018 and Seidman 2005 for a discussion on McDowell's "silencing" effect of virtue.

A change in moral outlook, thus, implies a change in what the agent sees as a possibility for action.⁹⁴

In short, one's moral outlook is the tool that allows the agent to make sense of and interact with the world in a morally significant way, and it is construed around some sort of foundations.⁹⁵ Imagine an agent fundamentally committed to loyalty: thanks to her moral outlook, she would be more sensitive than others towards occasions to be loyal and requirements of loyalty in each situation. In other words, just as the universe has been made intelligible through the Copernican paradigm, situations can be made intelligible through moral paradigms. Just as without any scientific paradigm we would not be able to even attempt to make sense of physical phenomena, without any moral paradigm we would not be able to navigate the world in a morally significant manner.

2.2 Scientific Paradigm Shifts and Moral Paradigm Shifts

So, there is such a thing as moral paradigms. Now, are there such things as moral paradigm *shifts*? Once again, I suggest we start from scientific paradigm shifts and see if we can draw a parallel in the moral domain. In what follows I will look at two narratives of scientific revolutions, the discovery of Uranus and Einstein's eureka moment, and argue that, in fact, the parallel stands. Thus, we have good reasons to think that there are such phenomena as moral change that comes about in a fashion that is

⁹⁴ Another possible framework is that of Taylor, who understands one's moral outlook as one's "moral map": in *Sources of the Self* (1989) he depicts the moral agent as the one whose commitments offer a sort of map to navigate the moral world; a moral world that is delimited by a horizon that is, itself, constitutive of human agency (1989: 27). For a different account of the "moral map" see Zagzebski 2017.

⁹⁵ An idea of a life worth living, in McDowell's case; commitments more broadly understood, in Taylor's case.

similar to that of scientific paradigm shifts, and which we can, therefore, legitimately call “moral paradigm shifts.”

A. Discovery of Uranus

“On at least seventeen different occasions between 1690 and 1781, a number of astronomers, including several of Europe’s most eminent observers, had seen a star in positions that we now suppose must have been occupied at the time by Uranus. One of the best observers in this group had actually seen the star on four successive nights in 1769 without noting the motion that could have suggested another identification. Herschel, when he first observed the same object twelve years later, did so with a much improved telescope of his own manufacture. As a result, he was able to notice an apparent disk-size that was at least unusual for stars. Something was awry, and he therefore postponed identification pending further scrutiny. That scrutiny disclosed Uranus’ motion among the stars, and Herschel therefore announced that he had seen a new comet! Only several months later, after fruitless attempts to fit the observed motion to a cometary orbit, did Lexell suggest that the orbit was probably planetary. When that suggestion was accepted, there were several fewer stars and one more planet in the world of the professional astronomer. A celestial body that had been observed off and on for almost a century was seen differently after 1781 because it could no longer be fitted to the perceptual categories (star or comet) provided by the paradigm that had previously prevailed” (Kuhn 1970: 189).

B. Einstein’s Eureka Moment

“The breakthrough came suddenly one day. I was sitting on a chair in my patent office in Bern. Suddenly a thought struck me: if a man falls freely, he would not feel his

weight. I was taken aback. This simple thought experiment made a deep impression on me. This led me to the theory of gravity” (Einstein 1922, translated by Yoshimasa A. Ono).⁹⁶

2.2.1 Passive

One obvious disanalogy between moral paradigm shifts and scientific paradigm shifts seems to be that scientists cannot but actively pursue discoveries of new paradigms, as opposed to being passive to them; that is, change does not happen to them in a passive way just like may be the case for moral change. But look at what happened to Einstein: what is more passive than sitting on a chair and letting imagination wander freely? Of course, if he hadn’t engaged extensively with his work, if he hadn’t possessed all the notions he did in fact possess, he would have not been able to reach such insight. Nonetheless, the insight happened in a moment of passive reception, rather than with active engagement with the subject matter. Before that insight, the new “ingenious description of the universe” (Knoblich & Oellinger 2006: 37) was not yet a possibility conceived by Einstein – until it was. Thus, at least (B) meets the *passive* criterion. However, I believe we have good reasons to believe that also (A) does meet the *passive* criterion, at least if we stick to Kuhn’s account of scientific revolutions.

In a recent discussion on Doctor Strange’s transformation from neurosurgeon to sorcerer, Shea (2018) argues that Kuhn’s scientific revolutions

“often looked quite a bit like Strange’s conversion experience, where scientists refused to change their fundamental ideas about how the world works, or their paradigms, until

⁹⁶ Notes from Einstein’s lecture in Tokyo in 1922. For a more romanticized version, see Knoblich & Oellinger 2006.

a series of crises forced them to. If and when they finally did change paradigms, they had to undergo their own version of Strange's conversion experience, which required them to relearn exactly what the "world" is" (2018: 1).

This seems to have implications for both the subjective and the objective/external level of paradigm shifts: on the one hand, we have scientists undergoing eureka moments;⁹⁷ on the other, we have a scientific community that, driven by single scientists' eureka moments, are forced to accept the new paradigm. In particular, what happens is that there are some scientists who are so deeply embedded in the business of "normal science" – i.e. puzzle-solving – under a particular paradigm, that they do not entertain any other conceptual possibility that may function as an alternative paradigm to the one they are currently using. This leads to the following situation: when confronted with what they cannot explain, scientists try to solve the puzzle by only looking at the conceptual framework that they have available. Any new possible conceptual framework is not a possibility. Note that here the claim is not that resistance to change is ideological, but just that, precisely because science is for the most part about puzzle-solving under a specific paradigm, when the puzzle seems impossible to solve given what is available, scientists have good reasons to believe that it seems impossible to solve *now*, but in the future the community at large may get there ("[Herschel] therefore postponed identification pending further scrutiny", example A). Then, every now and then, there is an "eureka!" moment, or, less dramatically, a moment in which a scientist looks at what everyone else is looking, and *sees* something different, which is what

⁹⁷ "Kuhn spends a great deal of time in *The Structure of Scientific Revolutions* thinking about why and how scientists make such revolutionary changes. He suggests that this process, at least at the level of the individual scientist, is both more mysterious and less prototypically rational, than scientists, historians, and philosophers have often thought." (Shea 2018: 2)

happened to Lexell seeing a planet instead of a strangely-behaving star or comet. This, in turn, forces them to re-adjust their entire conceptual framework:

“Adopting a new paradigm requires scientists to radically rethink not just their theories and methods, but also their overarching view of what the world is fundamentally like and what their role in this world is” (Shea 2018: 4).

Thus, there is a passive element that is present in both cases of scientific revolution: in (A) the scientific community is presented with a new paradigm that they did not entertain as a conceptual possibility until it was presented to them by the scientist who had the eureka moment; in (B) Einstein has the eureka moment, in which he is presented with a new paradigm that he did not entertain as a conceptual possibility until it was presented to him by his own passive daydreaming. Thus, scientific revolutions meet the passive parameter.

2.2.2 Abrupt

Kuhn (1970) draws a parallel with political revolutions to highlight an important factor of scientific revolutions: there is an element of subjectivity when it comes to perceiving revolutions as such. In his words: “scientific revolutions need seem revolutionary only to those whose paradigms are affected by them. To outsiders they may, like the Balkan revolutions of the early 20th century, seem normal parts of the developmental process” (93). Here, just as it was the case with our discussion of forms of moral change, there is a phenomenological difference between the subjective and the objective level that we need to consider. From a subjective perspective, scientists change the way they make sense of what they observe, and this can take the form of real “eureka!” moments; that is, of sudden, abrupt moments of disruption with how they had been looking at the same

phenomenon until a moment before. However, an external observer, say, a historian looking at example A, may pinpoint the moments in which paradigms have shifted, but would obviously not share the same phenomenology of those who brought such paradigm shift about. They may just observe an overall, gradual process of evolution with some points of disruption.

The element of experienced disruptiveness here is crucial. From the point of view of the scientist, changing rules and standards of practice in this revolutionary way has the following implication: scientists “see the world of their research-engagement differently. In so far as their only recourse to that world is through what they see and do, we may want to say that after a revolution scientists are responding to a different world” (111). In Gestalt terms:⁹⁸ “what were ducks in the scientists’ world before the revolution are rabbits afterwards” (ibid.). However, as we have seen before, Kuhn recognizes that this Gestalt shift does not automatically bring new means of research-engagement about: “at times of revolution, when the normal scientific tradition changes, the scientist’s perception of his environment must be re-educated – in some familiar situations he must learn to see a new gestalt” (112). Once the scientist has re-adjusted to the new paradigm, “the world of his research will seem, here and there, incommensurable with the one he had inhabited before” (ibid). He reinforces the point by employing the following example from psychological literature (Hanover Institute).⁹⁹ Inverting goggles possess lenses that flip images upside down and are used in psychological research to study

⁹⁸ Kuhn distinguishes these kinds of gestalt shifts from psychological ones, which, as opposed to scientific shifts, are “back-and-forth, with no ultimate authority to attest for perception” (1970/112).

⁹⁹ As quoted by Kuhn, see George M. Stratton, “Vision without Inversion of the Retinal Image”, *Psychological Review*, IV (1897), 341-60, 463-81. More up-to-date, still quoted by Kuhn: Harvey A. Carr, *An Introduction to Space Perception* (New York, 1935), 18-57.

image-building processes. A subject who wears inverted lenses goggles initially sees everything upside down:

“At the start his perceptual apparatus functions as it had been trained to function in the absence of the goggles, and the result is extreme disorientation, an acute personal crisis. But after the subject has begun to learn to deal with his new world, his entire visual field flips over, usually after an intervening period in which vision is simply confused. Thereafter, objects are again seen as they had been before the goggles were put on. The assimilation of a previously anomalous visual field has reacted upon and changed the field itself. Literally as well as metaphorically, the man accustomed to inverting lenses has undergone a revolutionary transformation of vision ...” (ibid).

Thus, from the perspective of the scientist who undergoes the eureka moment, there is a sudden, Gestalt-like change that is experienced as disruptive in relation to his previous perception. The change he undergoes is, therefore, *abrupt*.

2.2.3 Core

For moral values and commitments to be central to one’s paradigm, the following condition is necessary: if such values and commitments change, then the paradigm changes. As we have seen, a change in core values and commitments implies a fundamental change in moral outlook; that is, in the way agents interact with the world in a morally significant way. On the other hand, when values and commitments are merely peripheral, one can change them without necessarily bringing about a change in paradigm. The measure of centrality, thus, has to do with the scope of the readjustment that is needed in virtue of the change. In peripheral changes, one need not readjust to an entirely new framework of meaning, but possibly simply to small differences within that

very framework; in core changes, one does need to adjust to an entirely new framework of meaning, since the old one – the former paradigm – has been replaced.

Looking at A and B we can notice that, in both cases, postulates that were core to paradigms were replaced in a way that necessarily lead to the dismissal of former paradigms and, consequently, to a massive re-adjustment to the newly discovered ones. Former paradigms did not work anymore, they had to be changed because what was questioned and ultimately abandoned were the very core postulates – that is, the foundations of such paradigms.¹⁰⁰ In A, looking at Uranus with the star-paradigm or the comet-paradigm appeared obsolete, inaccurate, after the planet-paradigm was disclosed. There was no way to keep looking at Uranus with those paradigms, because the very foundations of those paradigms (I imagine, possible cosmic motions) underwent drastic reconfiguration. In B, looking at gravity with the Newtonian paradigm was not possible anymore: although the inaccuracy of the Newtonian paradigm was already accepted by the scientific community,¹⁰¹ there was still the possibility that the Newtonian paradigm could have been readjusted rather than replaced. But Einstein discovered that what had to be replaced were core conceptual frameworks of the Newtonian paradigm, namely, those that made sense of fundamental concepts such as space, time, matter, energy and gravity. Thus, there was just no way to readjust some superficial aspects of the Newtonian paradigm; it had to be replaced at once. In short, in both A and B the changes that came about necessarily lead to a change of paradigm. Thus, they were *core*.

¹⁰⁰ “In Kuhn’s terms, rival paradigms are *incommensurable*: there is simply no general method for translating the concepts taken from one paradigm into those of another. In a very real sense, scientists who adopt different paradigms literally experience different worlds, which contain very different sorts of things.” Shea 2018: 3. On Kuhn’s numerous accounts of incommensurability, see Sankey 1993.

¹⁰¹ Mercury’s orbit, for example, could not be fully explained by Newtonian physics.

We have seen that – at least some – cases of scientific paradigm shifts are passive, abrupt and core. Changes that are passive, abrupt and core, thus, can be classified as “paradigm shifts”. It follows that *moral* changes that are passive, abrupt and core can be legitimately classified as “*moral* paradigm shifts”.

The importance of MPS will be thoroughly considered in the next and final chapter of the dissertation. I will now take the remaining of this chapter to go back to our previous discussion on the three main conceptualizations of radical moral change – moral conversion, moral transformation and moral progress – and consider it once again in light of the proposed categorization of section 1. In particular, I am interested in finding instances of MPS in well-established narratives of radical moral change.

3. Looking for Moral Paradigm Shifts

As we have seen in the first chapter, moral changes that are radical and disruptive of what came before have already been considered by scholars, as well as by classical philosophers: let us see if we can spot MPS more clearly now that we have the help of the categorization of moral change that I proposed in section 1 of this chapter. Recall that the parameters are: (i) whether the change is actively pursued or passively received (A/P), (ii) whether they are experienced as abrupt or gradual (A/G), and (iii) whether they are core or peripheral (C/P), and that MPS are changes in the form of PAC: that is, they are passive, abrupt and core.

3.1 Moral Conversion

In Chapter I section 1 I proposed the following broad definition of Moral Conversion: “radical moral changes triggered by (more or less) sudden disclosure of value that leads to a variety of possible patterns of change”. Under this broad definition we can find conversions by evolution, by discovery, by revelation and epiphanic moral conversions. As proposed by Dees, his conversion by evolution seems to follow the PGC pattern; that is, it is passive, gradual and core. Passive to the extent to which he did not actively pursue it, gradual to the extent to which he himself explains it as a gradual development of ideas as opposed to an abrupt, disruptive change, and core, to the extent to which the change took place at a foundational level. What he calls conversion “by revelation”, on the other hand, seems to follow the Paradigm Shift pattern – i.e., the PAC pattern – since they are changes that are not actively pursued, that are experienced as abrupt to

the extent to which they do not naturally follow from developmental changes that came before, and take place at the core level. Moreover, they are reminiscent of religious conversion, since they are experienced as something “revealed” and passively received, just as it is the case for Malcolm X’s conversion to Islam.

What Dees calls conversion “by discovery” he exemplified by consciousness-raising groups. For the purpose of having a full picture on the matter, it is time to consider the phenomenon more in depth, since it has the particularity of belonging to individuals, but of also having political force. In her consciousness-raising manifest Kathie Sarachild writes that it started as a program organized by women who were radical and who wanted to study their own situation as women in society, to get to the root of issues, and not just to “take random action” (1973: 144). This is how she describes what happened in her group:

“One woman in the group, Ann Forer, spoke up: “ I think we have a lot more to do just in the area of raising our consciousness,” she said. “ Raising consciousness?” I wondered what she meant by that. I’d never heard it applied to women before. “ I’ve only begun thinking about women as an oppressed group,” she continued, “ and each day, I’m still learning more about it—my consciousness gets higher.” Now I didn’t consider that I had just started thinking about the oppression of women. In fact, I thought of myself as having done lots of thinking about it for quite a while, and lots of reading, too. But then Ann went on to give an example of something she’d noticed that turned out to be a deeper way of seeing it for me, too. “I think a lot about being attractive,” Ann said. “ People don’t find the real self of a woman attractive.” And then she went on to give some examples. And I just sat there listening to her describe all the false ways women have to act: playing dumb, always being agreeable, always being nice, not to mention what we had to do to our bodies with the clothes and shoes we

wore, the diets we had to go through, going blind not wearing glasses, all because men didn't find our real selves, our human freedom, our basic humanity "attractive." And I realized I still could learn a lot about how to understand and describe the particular oppression of women in ways that could reach other women in the way this had just reached me. The whole group was moved as I was, and we decided on the spot that what we needed—in the words Ann used—was to "raise our consciousness some more." (1973: 144-145).

This testimony does indeed share some elements of moral conversion, such as that of experiencing a – more or less sudden – disclosure of reality that calls for one's attention. In this sense, the "discovery" is that women are an oppressed group, both in society at large and in their familiar nucleus. However, this discovery is significantly different if compared to the one that Dees describes in his other two forms of moral conversion, "by evolution" and "by revelation." In consciousness-raising groups, the reality that is disclosed is not an individual reality, but a structural, political one.¹⁰² Dees writes that women realized that in order to change society they first had to change themselves (1996: 540), but this is a misinterpretation due to the fact that he sees that of consciousness-raising groups as an individual conversion, or a collection of individual conversions, instead of a phenomenon with collective force that cannot possibly be reduced to single women changing the way they go about in the world in order to change how they are treated. Of course, if we focus on Sarachild's testimony, we can appreciate the moment she discovered that there was much more to women's

¹⁰² "The personal is political," as stated by one of the most famous second-wave feminist slogans, see Hanisch (1970/2000). See also Rosen (2000) who explicitly reads Sarachild's consciousness-raising manifesto as the process by which women in small groups could explore the political aspects of the personal life" (2000: 197).

consciousness than she originally thought. That was, so to say, unexpected. But the whole group was originally moved by the wish to eradicate whatever it was that was making women's lives miserable: it was never about the individual women, but about women as a group.¹⁰³ Women did undergo a paradigm shift when it comes to framing their own situation, but unlike purely individual cases, where agents can, to some extent, re-structure their moral outlook accordingly, in this case there is not much *individual* women could do. The coexistence of individual and structural elements that is present in these type of cases makes it so that it is difficult to frame them simply as a collection of individual moral conversions.

Finally, Kristjánsson's epiphanic moral conversions are passive, in their being triggered by an encounter that is not actively pursued, and are definitely core, since they shake the foundations of one's moral apparatus. Less obviously can they be labeled as either abrupt or gradual: it rather depends on how the agent experiences such encounters. Thus, they can either fit the PAC – i.e. the Paradigm Shift pattern – or the PGC pattern.

We can conclude that Moral Conversions follow the Paradigm Shift pattern in the forms of a particular type of epiphanic moral conversions, a particular kind of conversions by discovery and quite straightforwardly in the form of conversions by revelations.

3.2 Moral Transformation

In Chapter 1 section 2 I isolated the following theoretical approaches to MT: MT as ascension and MT as activity. When it comes to MT as ascension, drawing a sharp

¹⁰³ Critiques to second-wave feminism include the fact that their concept of "oppression of women" did not consider the element of intersectionality, introduced and elaborated on by black feminist Kimberlé Williams Crenshaw (1989).

distinction between transformations that are actively pursued as opposed to passively received can be particularly challenging. It is true that one elevates one's self through contemplation, but it is not to elevate one's self that one engages in contemplation. Similarly, it is true that one reaches the form of truth at the end of TLP, but it is not to find the form of truth that one reads the TLP. Finally, it is true that one gains access to the truth if one pays loving attention to an other, but it is not to gain access to the truth that one pays loving attention to an other. The point of ascension is that one does not pay too much attention to one's self concerns: only then can one truly ascend. Furthermore, once the true form of reality is reached, one is in the position of merely receiving it, contemplating it, which is an obviously passive move. On the other hand, it is hard to deny that one engages in loving attention or reads the TLP in order to reach *something*, not just by accident. We might say that in order to undergo MT, one needs to be willing to take the spectator's shoes, just like the artist actively engages in bringing to life something that she will need to passively attend to, at some point of the creation process. This ambiguity when it comes to the active/passive distinction can be solved – although not exhausted – by sticking to one particular element of the active/passive distinction, which is that of actively pursuing the change or passively receiving it. In MT as ascension, one does not pursue the change *per se*; rather, one changes as a consequence of pursuing something higher – love, Beauty, the Good. Conversely, in actively pursuing something higher, one is in the position of passively receiving those radical changes that take place in virtue of contemplation of what is higher. Thus, if done right, MT as ascension is Passive to the extent to which one actively pursues the Higher, not because one wishes to change. Then, one also does undergo change, but only does so if one first lets reality guide one's transformation. Thus, only with a passive, contemplative attitude.

The gradual/abrupt distinction is a little easier to draw, since transformations are, by definition, gradual processes. The very fact that ascension is exemplified by the analogy of the ladder should suffice to classify it as a gradual, step-by-step process. On the other hand, there is a moment of “no going back”, which is when, according to Wittgenstein, we need to kick the ladder away. This may be read as a hint towards abruptness, since what characterizes abruptness is also an element of disruption: in gradual process you can always go back and make sense of what came before, while in abrupt processes what came before is not intelligible anymore, just like Malcolm X who could not understand why he was behaving the way he was before his conversion. In other words, we may think that after one has let reality reveal itself, one cannot just stop looking at it, or seeing the world through those new, realistic lenses. The Gestalt switch is somewhat irreversible. Again, these are nuances that depend on cases, and on the level of optimism we place in the possibility of truly reaching the truth. As we saw in Chapter I section 2, Murdoch would not be surprised seeing that no matter how much we transform, we always go back to step one, as if we kept climbing the ladder up and down. For the purpose of this research, we can say that when transformations take the shape of a discovery one cannot come back from, as it seems the case with the TLP, then they are abrupt to the extent to which they mark a discontinuity with what came before. When transformations take the shape of a contemplation of an other’s reality, as it is for Murdoch, then they may be either gradual or abrupt, depending on the particular case. That of M&D is likely to be a gradual change, both in terms of time and in terms of form of change, since M can always look back and make sense of why she was looking at D through her ego-centered concerns. Finally, MT as ascension take place at the core level, since they force the agent to adjust their moral outlook on whatever it is that they are contemplating. This also implies that the change is morally significant.

To sum up, MT as ascension is a passive, core process that can either be gradual or abrupt. When it is experienced as abrupt, it takes the form of a Moral Paradigm Shift.

When it comes to MT as an activity, we can see that the change patterns differ significantly from those of MT as ascension. To begin with, MT as an activity is less ambiguously active, since the agent decides to undergo the transformation with a clear goal in mind. Then, it can either be gradual or abrupt, depending on the particular case: if we look at Callard's recent account (2020), we can see that her example of learning ancient Greek takes the form of a gradual, step-by-step process, while Lenù's engagement in competitive activities with Lila leads to more abrupt changes, at least from her perspective. Finally, MT as an activity does not necessarily involve a transformation that is also morally significant. In this respect, we can see that some of Lenù's competitive activities transform her at the level of core values, while others do not. For example, she is drawn to question women's role in society and to engage in politics thanks to her blind emulation of Lila, which does transform her in a morally significant way – perhaps in a way that resembles the “moral conversion by discovery” of women in consciousness-raising groups, where the structural level is so entangled with the individual one that agent-level moral conversion would not itself exhaust the phenomenon of consciousness-raising. Nonetheless, in Chapter I we dwelled on the difference between engaging in a learning activity selflessly, à la Murdoch, and a learning activity that is goal-oriented, like Callard's aspirational competitions. The former involves a morally significant transformation, since in engaging in selfless activities one practices love. The latter may just lead to one reaching one's didactic goals, but nothing in that is necessarily morally significant.

Thus, MT as an activity can take the following patterns: AAC (active, abrupt and core), AGC (active, gradual and core), AAP (active, abrupt and peripheral) and AGP (active,

gradual and peripheral). What is significant for this research is that MT as an activity does not seem to follow the PAC pattern; thus, they do not take place as Paradigm Shifts.

3.3 Moral Progress

Finally, in Chapter I section 3 I have considered Moral Progress and its revolutionary dynamics. In particular, I have considered Individual Moral Progress as changes for the better in attitudes, behaviors and moral outlooks that come about in a radical manner; that is, as Gestalt switches, as discoveries of radically different ways to conceive of some state of affairs. At the individual level, the discovery that, say, what can conceive of animals as other living beings and not as food or utilities, is a discovery that is reminiscent of epiphanic moral conversions. I have also drawn a distinction between real and fake moral progress, where real moral progress is not just characterized by its being a move towards better states of affairs, but also and crucially by its revolutionary dynamics; that is, by its coming about as Gestalt switches in states of affairs perception. Fake moral progress, on the other hand, does not present this radical reconfiguration of outlooks. It can, thus, simply amount to development, or improvement of some sort.

In light of this, we can see that real moral progress comes about as abrupt, core changes that can either be actively pursued or passively received. That is, the Gestalt switch may come about in virtue of an active pursuit – just like scientists who actively look for better explanations of physical phenomena – or as a passive reception of an epiphany, as we saw in the analogy between moral paradigm shifts and scientific revolutions. In short, instances of real moral progress that are passive, follow the dynamics of moral paradigm shifts – that is, PAC. Of course, moral progress at the individual level can also

be gradual, thus follow the PGC or AGC patterns. What is interesting to note – going back to Wittgenstein’s critique – is that since moral progress, both at the societal and at the individual level, is paradigmatically characterized by its radical element of re-framing of outlook – the Gestalt switch – it is *always* core. In other words, a change that does not take place at the level of core values is not an instance of moral progress, since it is necessary that for real moral progress to be real moral progress, that the change amounts to a radical discovery, or reconfiguration of pre-existing paradigms. Otherwise, it would be just a mere scratching of fingers.

4. Conclusion

In this chapter I proposed a way to classify forms of moral change (section 1), depending on the following three elements: (i) whether the change is actively pursued or passively received, (ii) whether the change is gradual or abrupt, and (iii) whether the change is core or peripheral.

I then narrowed the scope of this dissertation to isolate one particular form of change, which is that of changes that are passively received, abrupt and core (PAC), which, in virtue of the analogy with Kuhn's paradigm shifts in science (section 2), I called Moral Paradigm Shifts (MPS). I then stressed the uniqueness of MPS' form of change, PAC, in relation to other slightly different triples. I considered AAC changes – that is, changes that are *active*, abrupt and core – where the change is actively pursued, or where the agent is at least in principle open to the possibility of change, just like our Marc going to group therapy. I also contrasted PAC with PGC – changes that are passive, *gradual* and core – and with PAP – changes that are passive, abrupt and *peripheral* – noting that when it comes to PGC, the difference is mainly phenomenological (i.e. whether the change is *experienced* as gradual or abrupt), while when it comes to PAP the difference concerns the fact that in peripheral changes, what it is that undergoes change is something morally relevant, but not morally significant. The example was Nora deciding to read a book instead of scrolling social media: it may indeed be better for Nora to read a book instead of scrolling social media, but this change in habits does not itself require that Nora re-shapes her entire value system. In short, it does not take place at a *core* level. Thus, notwithstanding the nuances, it still makes sense to draw distinctions among different forms of moral change.

In section 2 I drew a parallel between PAC changes and Kuhn's paradigm shifts, and argued that in virtue of the resemblance between PAC changes and Kuhn's paradigm

shifts in science, we can call PAC changes as “Moral Paradigm Shifts.” I started by stressing the fact that if scientists want to make sense of observed phenomena, they need a conceptual framework that defines the rules of meaningful engagement with such phenomena. Similarly, if agents want to make sense of moral situations, they need a conceptual framework that define the rules of meaningful engagement with such situations. Since scientific paradigms are necessary for scientific practice, then moral paradigms are necessary for moral practice. I then argued that we can make sense of moral paradigms through the notion of “moral outlook”, which is the tool that allows the agent to make sense of and interact with the world in a morally significant way, and that it is construed around some sort of foundations. I then concluded that there are such things as moral paradigms, and they can change. In the final part of section 2 I analyzed scientific paradigm shifts through the categorization proposed in section 1 and argued that they indeed follow the PAC change pattern: that is, they are passive, abrupt and core. Thus, it makes sense to talk of MPS as the moral version of PAC changes, both at the structural and at the individual level.

In section 3 I applied the classification of moral change to the three conceptualizations considered in Chapter I in order to look for MPS in the literature considered. There, I noticed that Moral Conversions follow the PAC pattern – and, thus, can be called MPS – when they are in the forms of epiphanic moral conversions, a particular kind of conversions by discovery and quite straightforwardly in the form of conversions by revelations. Concerning Moral Transformation as ascension, I noted that it is a passive, core process that can either be gradual or abrupt. Thus, when it is experienced as abrupt, it takes the form of a Moral Paradigm Shift. On the other hand, Moral Transformation as an activity can take the following patterns: AAC (active, abrupt and core), AGC (active, gradual and core), AAP (active, abrupt and peripheral) and AGP (active,

gradual and peripheral), but not the PAC pattern; thus, they do not take place as Paradigm Shifts.

Finally, I argued that real moral progress – that is, not the sort of change that amounts to mere growth – comes about as abrupt, core changes that can either be actively pursued or passively received. In short, instances of real moral progress that are passive, follow the dynamics of moral paradigm shifts – that is, PAC. I then noted that since moral progress is paradigmatically characterized by its radical element of re-framing of outlook – the Gestalt switch – it is *always* core. This implies that for moral progress to be real moral progress, the change needs to amount to a radical discovery, or reconfiguration of pre-existing paradigms. Otherwise, it would be just a mere scratching of fingers.

In the next and final chapter I will expand on this difference between mere growth and MPS. In particular, I will consider the possibility to account for MPS in a virtue ethical framework, given the inherent inhospitability of VE towards forms of change that do not amount to linear moral development, or mere moral growth, but that are, rather, disruptive of previous moral paradigms, just like MPS are.

References

- Bien, T. (2004). Quantum change and psychotherapy. *Journal of Clinical Psychology* , 35 (4), 493–501.
- Blasi, A. (1984). Moral identity: Its role in moral functioning. In W. M. Kurtines & J. J. Gewirtz (Eds.), *Morality, moral behavior and moral development* (pp. 128–139). New York: Wiley.
- Caprioglio Panizza, S., Vaccarezza, M. S. (2024). Moral Transformation as Shifting (Im)Possibilities. *The Journal of Ethics*:1-16.
- Dees, R. H. (1996). Moral conversions. *Philosophy and Phenomenological Research* , 56 (3), 531–550.
- Carr, H. A. (1935). *An Introduction to Space Perception*. Longmans, Green.
- Cederbaum, D. (1983). “Paradigms,” *Studies in History and Philosophy of Science*, 14(3), 173-213.
- Chappell, S. G. (2022). *Epiphanies: An Ethics of Experience*. Oxford: Oxford University Press.
- Crenshaw, K. W. (1989). ‘Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Politics.’ *University of Chicago Legal Forum*, 139–67.

Hanisch, C. (1970/2000). The Personal is Political. In *Radical Feminism*, ed. Crow.

Kohlberg, L. (1981). *The Philosophy of Moral Development Moral Stages and the Idea of Justice*. Harper & Row.

Kristjánsson, K. (2018). “Epiphanic Moral Conversions”, in *Self-Transcendence and Virtue*, Jennifer A. Frey and Candace Vogler (eds), Routledge, pp. 15-38.

Lapsley, D. K. (2008). Moral self-identity as the aim of education, in *Handbook of moral and character education*, Larry, P. Nucci & Darcia Narvaez (eds.), Routledge: New York.

MacLaughlin, A. J. (2008). *Narratives of hope: A philosophical study of moral conversions*. Ph.D. diss. Loyola University, Chicago.

McDowell, John. (1998) *Mind, Value, and Reality*, Harvard University Press, Cambridge.

Murdoch, I. (1970). *The sovereignty of good*. London: Routledge.

Knoblich, G., & Oellinger, M. (2006). The Eureka Moment. *Scientific American Mind*, 17(5), 38–43. <http://www.jstor.org/stable/24921587>

Rosen, R. (2000). *The world split open: How the modern women’s movement changed America*. New York: Viking Books.

Sanderse, W. (2015). An Aristotelian Model of Moral Development. *Journal of Philosophy of Education* 49 (3): 382-398.

Sankey, H. (1993). Kuhn's changing concept of incommensurability. *British Journal for the Philosophy of Science* 44 (4):759-774.

Sarachild, K. (1973). Notes from First National Conference of Stewardesses for Women's Rights in New York City, March 12,1973.

Shea, B. (2018). Paradigm Shift: A 'Strange' Case of a Scientific Revolution. In W. Irwin & White M. (eds.), *Dr. Strange and Philosophy: The Other Book of Forbidden Knowledge*. The Blackwell Series in Popular Culture and Philosophy. Wiley. pp. 139-150.

Siedman, J. (2005). "Two Sides of Silencing", in << *The Philosophical Quarterly* >>, 55, 218: 68-77.

Stratton, G. M. (1897). "Vision without Inversion of the Retinal Image", *Psychological Review*, IV (1897), 341-60, 463-81.

Taylor, C. (1982). The diversity of goods. In Amartya Sen & Bernard Williams (eds.), *Utilitarianism and Beyond*. New York: Cambridge University Press. pp. 129--135.

Taylor, C. (1989). *Sources of the self: The making of the modern identity*.

Vigani, D. (2019). “Virtuous Construal: In Defense of Silencing”, in *Journal of the American Philosophical Association*, 5, 2: 229-245.

Walden, K. (2015). Art and moral revolution. *The Journal of Aesthetics and Art Criticism* 73: 283–95.

Zagzebski, L. Trinkaus. (2017). *Exemplarist Moral Theory*. New York, NY: Oup Usa.

Chapter III. Virtue Ethics and Moral Paradigm Shifts¹⁰⁴

This chapter has the following purposes:

1. Explore how MPS impact VE as a family of moral theories (in particular, neo-Aristotelian VE);
2. Isolate VE's resistance to MPS;
3. Propose a way for VE to facilitate MPS.

In this respect, I will advance the following arguments:

Argument 1.

- (i). MPS pose a “flexibility requirement” on moral theories (from Chapter II);
 - (ii). If VE can provide a flexible account of virtue acquisition, then VE meets the MPS requirement;
 - (iii). VE does not provide a flexible account of virtue acquisition;
- => VE does not meet the MPS requirement.

Argument 2.

- (i) if VE shifts the focus on epistemic virtues, rather than moral virtues, then VE can be a flexible moral theory;
- (ii) if VE rests on a non-realist notion of “flourishing”, VE can provide a dynamic account of flourishing;
- (iii) if VE can in principle be flexible, then it should be so;
- (iv) VE can shift the focus on epistemic virtues (virtue epistemology);

¹⁰⁴ This chapter is an expansion of the following co-authored paper under review: “Individual Moral Progress. A virtue-based approach”, with F. Bina, M. Reichlin, S. Songhorian and M. S. Vaccarezza.

(v) VE can rest on a non-realist notion of “flourishing” (ethical naturalism without metaethical naturalism);

=> VE can in principle be flexible.

=> VE should be flexible.

In Chapter I I explored radical moral change in its three versions of moral conversion, moral transformation and moral progress. In Chapter II I proposed a possible categorization that we could use to make sense of forms of moral change that revolves around three parameters: (i) whether the change is actively pursued or passively received (A/P), (ii) whether the change is experienced as gradual or abrupt (G/A), and (iii) whether it is core or peripheral (C/P). I then isolated one specific form of moral change, which is that of PAC (passive, abrupt and core), and argued that this form of moral change has the potential to disrupt and replace existing moral paradigms with new, more accurate ones. For this reason, they can be called “Moral Paradigm Shifts” (MPS).

In this chapter I will consider MPS within the moral framework I am mostly familiar with, which is that of Virtue Ethics. My own familiarity with the topic does not exhaust the choice: Virtue Ethics as a family of theories considers the development of virtues as central to the formation of morally decent individuals. Thus, Virtue Ethics as a family of theories conceives of developmental moral change as central to the moral formation of individuals. My interest is that of exploring whether the form of change that is core to Virtue Ethics can be a form of MPS and, if not, whether there can be room at all for MPS in a Virtue Ethical framework. The answer to the first question is, quite obviously, no: moral development is not a PAC change, thus it is not in and of itself a form of

change that is akin to MPS. The answer to the second question, at first sight, is also no: there seems to be an inherent resistance of Virtue Ethical accounts to forms of moral change that are disruptive of existing moral paradigms. However, it would be important for Virtue Ethical accounts to make room for the possibility of transcending one's moral paradigm, and not just develop according to existing ones, as it is the case for linear moral development.

In what follows, I will isolate the problem of conservatism that seems to be inherent to Virtue Ethical methodology and attempt to draw a possible way out within a Virtue Ethical framework. It is important to note that "Virtue Ethics" is way too broad of a notion to be used fruitfully. As mentioned before, it is a "family of theories" (on which I will expand in section 1) that needs to be narrowed down to specific Virtue Ethical accounts in order for its use to be meaningful in any way. In this dissertation I focus on neo-Aristotelian Virtue Ethics; that is, on the version of Virtue Ethics that considers virtues as dispositions of character, flourishing as the end of moral action, and the virtuous act to do as something "natural" (Snow 2017).

I begin by providing an overview of neo-Aristotelian moral development (section 1.1), and stressing the fact that it would be important for neo-Aristotelian VE to be at least in principle equipped to make room for the possibility for individuals to transcend their moral development through MPS (section 1.2). In this regard, I present the conservative charge against VE (section 1.3) that highlights the fact that VE is inherently conservative in methodology and, thus, unequipped to make room for changes that are *progressive*; that is, that are paradigm-disruptive rather than paradigm-reinforcing.

I then ask whether it would be possible to conceive of virtues in a non-conservative way. To this end, I explore two main ways to conceive of virtues: the first one sees virtues as arising out of moral traditions (section 2.1), the second one sees virtues as arising out of

facts about human nature (section 2.2). I stress that both these ways to make sense of virtues present problems when it comes to making room for progressive moral change of the MPS type. I end the section with the following argument: making sense of virtues in terms of moral traditions is a conservative methodological move, but making sense of virtues as arising out of nature need not be. Second-nature naturalism à la McDowell does not restrict human nature to scientific, inflexible facts about human nature, which is a promising way to give neo-Aristotelian VE an objective action-guidance, without necessarily doom it to conservative, inflexible accounts of human nature. On the other hand, it is not clear how second-nature naturalism successfully accounts for nature's normativity. Swanton (2021) has recently tried to meet this problem by giving metaphysical grounds to McDowell's second-nature naturalism, ending up with the following picture: second nature is normative in virtue of the fact that we exist as entities in a moral network of meaning, where "thick" virtue concepts make sense because by giving meaning to what there exists, they also exist in the world. As Beings-in-the-world, we can see the requirement of virtue as a meaningful way to interact with other Beings-in-the-world. This, according to Swanton, solves McDowell's problem of being able to account for the inherent normativity of what one *sees* in the situation. I argue that metaphysical solutions are both unnecessary and counter-productive, since they make naturalistic claims all the more inflexible by grounding them not only in facts, but also in metaphysical entities. I end the section with a suggestion on how to make sense of naturalism without metaethical naturalism, in a way that both maintains its normative force, thus maintaining the neo-Aristotelian naturalistic element, and does not preclude the very notion of "human nature" and, relatedly, of "flourishing" to be a rigid, inflexible, hence conservative theoretical pillar of neo-Aristotelian VE.

In the third and last section I argue that if neo-Aristotelian VE can integrate moral development with epistemic development, then it could be equipped to make room for the possibility of MPS and, thus, be a progressive rather than a conservative moral theory. This is because a moral development that is also epistemic would allow agents to develop a moral outlook that is at least in principle equipped to undergo MPS and readjust their moral outlook accordingly. In this picture, neo-Aristotelian normative elements such as “flourishing” and “human nature” would also be subject to radical moral change throughout agents’ lives; something that, as of now, is still not conceived as a possibility. I then propose a by no means exhaustive list of epistemic virtues that, if developed, could prove to be “transformative” to the extent to which they would go to construe a moral outlook that is capable of dealing with MPS. I conclude with two practical examples of how a progressive moral outlook would look like and would work through radical moral transformations, whenever time requires that one adjusts to moral complexities that are totally new and in need to be taken seriously.

1. Moral Change and VE

MPS require moral theories to be flexible. Let us unpack this: if it is true that people undergo MPS, then moral theories that prescribe what is the right thing to do and what should guide our actions need to allow for some flexibility of what the right thing to do might be and should guide our actions. This is not to say that we should not believe in moral theories’ prescriptions, nor that there is nothing ultimately right or good (moral relativism). It just means that what we deem right now may be subject to radical moral change for the better: that is, even if we have good reasons to deem something as right now, it does not rule out that there will be a better way to make sense of the right thing

to do in the future. Just like scientists had really good reasons to follow Newtonian physics before Einstein's paradigm shift, we surely have really good reasons to regard certain things as right and others as wrong now.

Developmental theories such as neo-Aristotelian VE revolve around the development of virtues according to some fairly inflexible pillars: virtue acquisition and human flourishing. If this is the case, then VE is destined to be an obsolete moral theory: that is, one that allows agents to morally develop, but not to morally progress. We will see that flexibility in this sense would not undermine prescriptions, thus maintaining the status or moral theory. Let us take a step back.

Virtue Ethics is generally considered one of the three main normative theories, along with Consequentialism and Deontology. It is a family of theories that conceives of virtues as central for moral action. VE accounts differ concerning how they define virtues; in particular, concerning how they make sense of *phronesis*, which is generally deemed as the fundamental virtue that allows one to perceive when, how and to what extent one should, say, be honest or charitable.

We can find Eudaimonistic VE, Agent-based VE, Platonic VE and, recently revived, Target-centered VE. Eudaimonistic VE conceive of virtues in relation to what the life of eudaimonia may amount to (LeBar 2013): that is, the good life is the eudaimon life, and what allows one to live the eudaimon life is developing virtues. Most VE accounts link developing virtues with a life of eudaimonia, but not all of them would argue that developing virtues is both necessary and sufficient for a life of eudaimonia. Agent-based accounts of VE, for example, stress the importance of agents' moral attributes, such as their motivational and dispositional qualities (Slote 1997, 2001) when it comes to distinguishing between right and wrong actions. Right actions are those that come

about in virtue of having the right motives, while bad actions come about in virtue of bad motives. Zagzebski (2004) reads this differently: what provides normative standards are the motivations and dispositions of exemplary virtuous individuals.¹⁰⁵ What eudaimonia amounts to, in this case, is also subject to these sorts of exemplarist evaluations. Platonic accounts of VE, on the other hand, do not define goodness in terms of virtue: the Good is a different metaphysical entity and virtues are what human beings can develop in order to draw close to it (Chappell 2014; Murdoch 1971).¹⁰⁶ Finally, according to Swanton's Target-centered VE (2003, 2021), virtues have a field and a target: virtuous acts are those that, given the particular field of the virtue, successfully hit the target. For example, generosity's field is that of sharing resources; a generous act is one that successfully hits the target of sharing resources. This account takes the focus away from the agent, as well as from eudaimonia, since what the target of each virtue is, and how to successfully hit it, depend also on factors that fall outside the agent's control, and since hitting the target not necessarily make one happy (Swanton: 2021).

As we can notice, that of VE is a family of theories that can be so different among each other that it makes it hard to speak of one, ultimate "VE theory," or "VE account." What is important to note is that all VE accounts share the fact that virtues are fundamental to doing the right thing, or developing a virtuous character, or living a flourishing life. Swanton calls it "The Centrality of Virtuousness" assumption (2021: 5) that is found in all VE accounts. In other words, *developing* virtues is fundamental to

¹⁰⁵ It is worth noting that some virtue ethicists would have no problems claiming that VE has no business in telling what the right thing to do is, but only what kind of person we should be: see MacIntyre (1985) and Anscombe (1958). More recently, Brewer (2009).

¹⁰⁶ As we have seen in the first chapter, Iris Murdoch shares the Platonic metaphysics, while it is less clear whether she was also a virtue ethicist.

doing the virtuous thing. Thus, linear, step-by-step changes towards the goal of becoming virtuous, or developing virtuous dispositions, is the sort of moral change that the VE family seems most interested in. On the other hand, VE accounts hardly ever mention anything akin to MPS as a particularly noteworthy dynamic of moral change.

In what follows I will focus specifically on neo-Aristotelian accounts of VE: they are the ones most interested in the dynamics of change when it comes to developing virtues, thus the ones more likely to be able to integrate the possibility of MPS in their spectrum of moral change.

1.1 Neo-Aristotelian moral development

According to the neo-Aristotelian model, virtue acquisition is a gradual, step-by-step process towards moral maturity, which is what guarantees that the virtuous thing to do is brought about in each situation (Carr and Steuel 1999; Kristjánsson 2007, 2015; Sanderse 2015). This is hardly what Aristotle had in mind, since he understood the aforementioned types of moral characters – indifference, *akrasia*, *enkrateia* and virtue – as categories, rather than as stages that one could climb. Nonetheless, neo-Aristotelians claim non only that it is possible to understand Aristotle’s categories as developmental stages, but also that it is both very much in line with Aristotle’s naturalistic view of development¹⁰⁷ and advantageous if we wish to understand developmental moral psychology and education, since it helps us understand the constitution of moral character (Sanderse 2015). According to the Aristotelian stage model, “stages” are

¹⁰⁷ there are controversies: Aristotle would not have claimed that children are potentially virtuous (NE 1111b4-10; Kristjánsson, 2007, p. 22; Sherman, 1989, p. 161). But, argues Sanderse, “insisting a gap exists between children and adults in terms of morality and rationality is compatible with the idea that children can learn to bridge the gap.” (Sanderse 2015: 5; *Politics* 1260a32-33; Sherman, 1989, p. 162).

“ ‘progressive levels of moral excellence’ that correspond ‘to different developmental conditions of an agent’s soul’ ” (Kristjánsson, 2007, p. 20; Silverstein and Trombetti, 2013, p. 237 in Sanderse 2015: 6).¹⁰⁸ In other words,

A stage is a level of alignment between reason, desire and action and the sort of pleasure that accompanies it. The configuration of these components gradually changes as a person reaches a new stage. This can result in a complete metamorphosis: the virtuous person experiences and interprets the world differently than the morally indifferent person (Sanderse 2015: 6)

The stage of moral indifference is characterized by a kind of happiness that coincides with “low” pleasures bodily pleasures, honor and money (NE 1095a22-23) and by a lack of habituation to what is truly fine and pleasant; thus, people at this stage do not have the ability to determine what the virtuous thing to do is and why (Sanderse 2015: 7).¹⁰⁹ The second stage, that of *akrasia*, corresponds to the stage in which people have “chosen to lead a virtuous life, instead of a life guided by the desire for wealth or fame” (Sanderse 2015: 8). They see the attractiveness of a life of virtue, yet they lack the ability to control their passions; that is, they lack the ability to direct their passions

¹⁰⁸ Not all Aristotelians agree with this number of identifiable stages. For example, Sherman (1989) and Tobin (1989) distinguish between three stages, the ‘novice’ or ‘initial’ stage, an ‘intermediate’ stage and finally the stage of ‘full virtue and rationality’. Garrett (1993) claims that Aristotle’s ladder has five steps: the ‘fully wretched’, ‘most other people’, the ‘incontinent’, the ‘continent’ and the ‘virtuous’. Similarly, Curzer (1998, 2002, 2012) distinguishes between the ‘incontinent’, ‘continent’ and the ‘virtuous’, but he also introduces the ‘generous-minded’. See also Kristjánsson (2007, 2010), who adds the levels of the ‘soft’ (or ‘self-indulgent’) and the ‘resistant’.

¹⁰⁹ See NE 1095a22-24, NE 1142b30, NE 1116b30-35.

towards virtuous goals.¹¹⁰ Agents develop from the first to the second stage when they have acquired the ability to discern between vicious and virtuous acts.¹¹¹ When one manages to develop self-control and, thus, manages to redirect her passions towards virtuous goals, she has entered the stage of *enkrateia*: she acts in accordance to virtue, but acting virtuously still does not come naturally.¹¹² Finally, the last stage is that of proper virtue, which Aristotle contrasts with natural virtue (NE 1144b3): the one who is gifted with moral qualities by nature merely resembles the properly virtuous, since what the naturally gifted one lacks is *phronesis*, which is the intellectual virtue that guarantees that the virtues are forged into a “complete, stable and integrated collection” (Sanderse 2015: 12). Thus, the truly virtuous is not only occasionally virtuous, or virtuous only when it comes to one particular sphere (Curzer 2012). Rather, “[v]irtuous people [...] have a stable and firm commitment to the good over a lifetime and hit the mean with regard to actions and emotions in all spheres” (*ibid.*). They also derive joy from acting virtuously, even if their overall quality of life may not be better off than those of other stages. Interestingly, Sanderse stresses the fact that this last stage is open-ended, to the extent there is always room for improvement, for example in how to act virtuously, or how to understand virtue in the first place (2015: 12). Thus, practically speaking, the end-stage of moral development cannot be fixed in advanced (Annas 2011).

¹¹⁰ See NE 1147a24, NE1150a18-21.

¹¹¹ Generally, through shame (Sanderse 2015: 9).

¹¹² That is, as a “second nature.” That of the difference between *enkrateia* and proper virtue is a tricky distinction; one may even wonder whether it makes sense to draw it sharply (Sanderse 2015: 10). I have argued elsewhere that we can draw a sharp cognitive distinction between the *enkrates* and the fully virtuous (Liberti 2022).

According to Sanderse, the Aristotelian stage model offers philosophers a useful tool to understand what happens to people in different stages of moral development, since it offers psychologically accurate descriptions of diverse and realistic character types. Moreover, it offers moral educators a useful and accurate model around which to structure educational interventions that are aimed at reaching the final stage of proper virtue.

Although it is indeed true that this fairly innovative account of Aristotelian moral development as an open-ended and dynamic process has a lot to offer when it comes to moral change, there is still no room for what I have presented as “moral paradigm shifts”. That is: in Aristotelian moral development agents develop in accordance with an existing moral paradigm and develop into fully virtuous agents in accordance with that moral paradigm. It is, however, important to be able to account for possible changes in moral paradigm; specifically, it is crucial for moral education to elicit the capacity for potential disruptions of paradigms. In the next section, I will expand on the reason why it would indeed be required for a moral theory to be able to account for MPS.

1.2 The need for Moral Paradigm Shifts

In a recent paper Rehen and Sauer (2024) have highlighted the fact that moral education, contrary to what some scholars have thought,¹¹³ serves moral traditions and is, thus, not

¹¹³ See for example Morrow (2020, p. 503) who first argues that “plausible theories of moral progress must provide an account of the pathways and processes by which progress occurs” then considers “moral education, achieved through exposure to compelling texts, images, or oral narratives” as an instance of such process. Also Hermann (2017, p. 46) agrees: “the education of capacities and virtues [...] can contribute to the achievement of moral progress.”, as well as Schinkel and de Ruyter (2017, p. 121): “that moral progress is possible is a foundational assumption of moral education.”

a good candidate when it comes to enabling moral progress. Theirs is not much of a critique; rather, in the paper they provide empirical evidence to support the claim that, generally, moral education reinforces traditions, be it a good or a bad thing; thus, if we wish to look for enablers of moral progress, we should look elsewhere. The key issue is that moral education provides stability, or continuity: that is, stability or continuity with an existing moral paradigm. Thus, moral education does not *cause* moral progress; at least, not often (2024: 26). They note that, especially in virtue of this element of stability, moral education may, on the other hand, prove useful when it comes to stabilize progressive achievements: this is, when the change of paradigm has taken place and people need to adjust to it (ibid.).

Their study focuses on moral education as a form of social learning. They define “moral education” as “any activity explicitly aimed at or geared towards improving the moral beliefs, values or behaviors of children and adolescents” (2024: 26) and “social learning” as “learning that results in the social transmission of information, which occurs when one individual who has acquired a piece of information (or a behavior), expresses that information (for example, says the information out loud, or behaves in a way that depends on the information), and this exerts a lasting causal influence on the rate at which another individual acquires the same information” (ibid.). Character education fits this description, since it comprises of social activities that are explicitly aimed at improving people’s moral capacities;¹¹⁴ thus, the findings of Rehen and Sauer’s paper are also relevant for character education.

To develop their study, they start from the following assumptions about moral education:

1. “the moral information that educators attempt to transmit mirrors the moral beliefs, rules, values, etc. of their group” (2024: 30).

¹¹⁴ See for example Kristjánsson 2015 and The Jubilee Center for Character Education in Schools (2022).

2. “all educators reach similar proportions of the pool of potential learners; no one educator reaches, for example, ten times more learners than every other moral educator” (ibid.)

3. “old and new moral information have the same likelihood of being transmitted” (ibid.)

According to assumption (1) moral educators attempt to transmit information that mirrors existing moral paradigms. Moreover, (2) all moral educators have pretty much the same level of influence, and (3) if they attempt to transmit moral information that is incompatible with the existing moral paradigm – and, thus, potentially paradigm-disruptive – the moral information they attempt to transmit has the same likelihood of being transmitted as moral information that is compatible with the existing moral paradigm. When all these assumptions are true, moral education does not play any major role when it comes to bringing about moral change, be it progressive or not. Rather, when all these assumptions are true, moral education functions “as a force of moral stability or continuity” (2024: 31).

Then, they evaluate their assumptions in the face of instances that could violate them. For example, assumption (1) can be violated when large institutions (e.g. church or state) that dictate moral curricula in schools are particularly progressive (which is generally not the case), or when teachers with moral beliefs that are radically different from the status quo have also a significant say when it comes to transmitting their moral information (2024: 33). Assumption (2) is, according to the authors, harder to violate, since “time and biology both constrain the number of children that any individual parent can attempt to educate within their lifetime, so it will be rare for any one parent to have an out-sized influence on the population of prospective moral learners. The same strikes us as reasonable for school-based moral education, on similar grounds” (2024: 33). However, they leave the possibility for status-bias to violate this assumption when the feeling of emulation towards specific individuals is so widespread that it can reach out-

sized influence.¹¹⁵ Finally, assumption (3) can be violated by the presence of content-bias, where some content is transmitted more successfully than another and this depends on the content itself. Content-bias is, note the authors, very much linked to emotion bias, where “[t]he idea is that information that elicits an emotional reaction from the learner is more likely to be transmitted compared to information that does not” (2024: 35).¹¹⁶ While emotions may be a good way to push for progress (Appiah 2010; Hu, 2021; Sauer, 2019), there is no guarantee that emotional bias would systematically lead to changes that are progressive.

Thus, Rehen and Sauer conclude that “[m]oral education can be a force for change when moral educators transmit information that differs from the moral beliefs and values of their society, when a small group of like-minded educators are disproportionately influential, or when different pieces of moral information transmit at different rates” (2024: 37). However, they stress that it is one thing to contribute to moral change, quite another to contribute to *progressive* moral change, and whether moral education does contribute to progressive moral change “will depend on the details of each episode of moral change” (ibid.).

We can draw the following conclusion: given that VE stresses the education of moral character as its core feature when it comes to developing virtues, then for VE to be progressive it needs to account for the possibility that the education of moral character can itself be transcended. Only if this is the case, can VE theories be progressive, in their not functioning merely as status quo reinforcing developmental theories concerning virtues. In the next chapter, we will see that the question of the inherent conservatism of VE theories can be both theorized into a proper objection to VE, and fully met by VE theories that wish to do so.

¹¹⁵ See Croce, Liberti and Vaccarezza 2024 on the power of moral influence.

¹¹⁶ See Eriksson & Coultas 2014; for an experiment, see Nichols 2002.

1.3 The problem of Conservatism

The Stanford Encyclopedia of Philosophy defines conservatism as a “social attitude” that “expresses the instinctive human fear of sudden change, and tendency to habitual action” (Hamilton 2020). A bit more specifically, one can draw a distinction between conservatism as an attitude and conservatism as a stance, or what Beiser (1992) called “self-conscious” conservatism, where conservatism develops as a resistance in opposition to movements such as the Enlightenment or the French Revolution. In general, Conservatism has a preference for what has already worked in the past.¹¹⁷ Thus, conservatism reserves a central role to tradition.¹¹⁸ When it comes to the dynamics between traditions and new, pressing issues, conservatives are generally reformists: that is, traditions may be reformed, adjusted to current changes and needs, in order to maintain their action-guidance. Thus, according to conservatives, what has worked in the past *can* be reformed to meet the needs of the present; but this is still a linear change where the overall traditional paradigm is not replaced by a new, more accurate one. Replacement of traditional paradigms is not at issue.

In a recent paper, Jakob Ohlhorst (2023) argues that virtue ethics is inherently conservative, since it reserves a central role to tradition and it aims at conceptual reforms rather than conceptual revolutions. For this reason, virtue ethics is particularly unequipped when it comes to making sense of, as well as making room for, moral progress. Specifically, Ohlhorst stresses how virtue ethics methodologically “conceives of virtues in the same terms and with the same perspective as its ancient and mediaeval precursors as well as our common sense conception” (1); and it is this appeal to

¹¹⁷ As noted in the entry, both Aristotle and Confucius are generally taken as the precursors of self-conscious conservatism (Hamilton 2020).

¹¹⁸ Although the contraposition between tradition and modernity becomes explicit in the 18th century (“the invention of tradition”, Hobsbawm and Ranger 1983).

common sense conception of virtues that makes virtue ethical methodology so resistant to progress. He notes that

The default philosophical method in virtue ethics is conceptual analysis. That is, usually, virtue ethicists try to discover the true nature of particular virtues by analyzing the concepts that designate these virtues. This is the same project that Plato pursued in the *Republic*: Answering the question “what is justice?”. This approach is based on our intuitions about what is and what is not designated by particular virtue concepts. (2)

Scholars working in VE typically take as their starting point what we already think of virtues, and then try to accommodate such knowledge and intuitions to fit current issues; the most that this methodology can do is to reform pre-existing understandings of virtues. In fact, Ohlhorst notes that “[c]onceptual analysis is a conservative method in the sense that it is only revisionary, i.e. rejects tenets and conceptions, when it cannot be helped due to internal contradictions of the common intuitions or with our scientific knowledge and the goal of such a revision is only to purify the concept to be analysed from these contradictions” (4). In short, virtue ethics focuses on what we already think rather than on what we should be thinking, independently on how we have understood it so far.¹¹⁹ This suggests that one way to get around the problem of conservative would be to search for the possibility for VE theories to make sense of virtues in a way that does not necessarily tie them to inflexible parameters. That is, if in order to make sense of virtues we could rely on new conceptual devices, rather than on old, or inherently rigid ones, then one would not need to merely revise virtue concepts; virtue concepts would themselves be open to the possibility of being framed in a radically innovative

¹¹⁹ According to Ohlhorst, agent-based virtue ethics like Zagzebski’s (1998; 2004) or Slote’s (1995) offer prime illustrations of this conservative method.

way. In what follows, we will explore the conceptual origins of virtues, and see that, in fact, they present problems to the possibility of accounting for virtue concepts in a progressive manner.

2. The (In)flexibility of Virtues

Here we will see that VE does not provide a flexible account of virtue acquisition. Thus, VE fails to meet the MPS requirement. This is because VE sees virtues as arising either out of tradition (MacIntyre), or as deriving from facts about human nature (Foot). Traditions are not flexible, and neither is Foot's naturalism. However, neo-Aristotelian naturalism is also what provides the normative element of VE theories. I will thus argue that, to begin with, VE needs an account of naturalism that can both maintain its normative element ("this is what the good looks like, therefore this is what we should strive for") and is flexible.

Where do virtues arise from? Or, relatedly, on which basis can we justify a list of the virtues we need to develop? When it comes to these questions, there are generally two trends in VE accounts: one, which has most famously advanced by MacIntyre replies "in tradition," while the other, whose most prominent contemporary proponent is Foot, replies "in facts about nature". Both these answers are problematic, since neither seem to consider the fact that traditions and normative conceptions of human nature *should* be subject to MPS for there to be progress. To illustrate, the way we understand the virtue of "modesty" should be subject to MPS if we wish to liberate women from judgments on appearance; likewise if one considers rape something "natural" because some species of animals show rape-like behavior should be subject to MPS.

In what follows I will first consider MacIntyre's account that sees virtues as arising from traditions; then in 2.2 I will consider Foot's alternative account on the origins of virtues, and highlight in what ways they are both inhospitable to the possibility of MPS.

2.1 VE and Tradition

As argued in Part II, moral paradigms exist and are subject to – more or less radical – change. In *After Virtue* (1985) MacIntyre proposes an analysis of different “paradigms”, or traditions, of morality: the Homeric, the Aristotelian, the New Testamental, that of Jane Austen and of Benjamin Franklin. He begins by contrasting the first two and notices that

[...] it is not just the case that the difference between Aristotle and Homer lies in the inclusion of some items and the omission of others in their respective catalogues. It turns out also in the way in which those catalogues are ordered, in which items are ranked as relatively central to human excellence and which marginal. Moreover the relationship of virtues to the social order has changed. For Homer the paradigm of human excellence is the warrior; for Aristotle it is the Athenian gentleman. Indeed according to Aristotle certain virtues are only available to those of great riches and of high social status; there are virtues which are unavailable to the poor man, even if he is a free man (1985: 182).

So, it is not just that in these two different paradigms the virtues were different, but also – and more interestingly – that those that were common to both paradigms were ranked differently. What is of central importance in the Homeric paradigm was not of central

importance in the Aristotelian paradigm. Accordingly, paradigmatic exemplars change: for Homer the paradigmatic virtuous agent was the warrior, while for Aristotle was the gentleman. MacIntyre goes on to remark that “the most striking contrast with Aristotle's catalogue is to be found neither in Homer's nor in our own, but in the New Testament's” (ibid.), since

the New Testament not only praises virtues of which Aristotle knows nothing – faith, hope and love – and says nothing about virtues such as *phronesis* which are crucial for Aristotle, but it praises at least one quality as a virtue which Aristotle seems to count as one of the vices relative to magnanimity, namely humility. Moreover since the New Testament quite dearly sees the rich as destined for the pains of Hell, it is clear that the key virtues cannot be available to them; yet they *are* available to slaves. And the New Testament of course differs from both Homer and Aristotle not only in the items included in its catalogue, but once again in its rank ordering of the virtues (ibid.).

Once again, we can observe how different moral traditions can look at the same phenomena and classify it in two radically different ways. Take being rich: according to the Christian paradigm of the New Testament, being rich is something one should be ashamed of, since it quite automatically elicits the judgment that one's soul is corrupted; that is, not destined to Heaven. On the other hand, if we could inhabit the Aristotelian paradigm, we would take wealth as the prerequisite of being virtuous, not as an obstacle to it. Just like in a Gestalt switch, a change in moral tradition impacts agents' moral outlooks, to the point that they look at the same phenomenon that has been observed by previous traditions and see something that is completely different. Now, MacIntyre's questions do not concern the existence of and the relation between moral paradigms: he is concerned with the possibility of an objective account of what virtues are; that is, he

is looking for the common grounds, the substance of virtue that is present in all accounts he considers. In this respect, he notices that

[o]ne of the features of the concept of a virtue which has emerged with some clarity from the argument so far is that it always requires for its application the acceptance for some prior account of certain features of social and moral life in terms of which it has to be defined and explained. So in the Homeric account the concept of a virtue is secondary to that of *a social role*, in Aristotle's account it is secondary to that of *the good life for man* conceived as the *telos* of human action and in Franklin's much later account it is secondary to that of utility (186).

In other words, in all traditions of virtue, what virtue is comes as secondary; it has to be defined in terms of something else. More specifically, what it means for something to be “good” or “bad” is determined by specific “practices”, which MacIntyre defines as

any coherent and complex form of socially established cooperative human activity through which goods internal to that form of activity are realized in the course of trying to achieve those standards of excellence which are appropriate to, and partially definitive of, that form of activity, with the result that human powers to achieve excellence, and human conceptions of the ends and goods involved, are systematically extended (187).

In this sense, moral paradigms are a bit like games. There are certain rules that set standards of excellence – e.g. how to checkmate your opponent – and that, thus, define what is good or bad, appropriate or inappropriate for that kind of practice. It is in this

sense that “good is internal to practice” (188):¹²⁰ entering a practice implies that I accept the “authority of the best standards realized so far” (190). Thus,

if, on starting to play baseball, I do not accept that others know better than I when to throw a fast ball and when not, I will never learn to appreciate good pitching let alone to pitch. In the realm of practices the authority of both goods and standards operates in such a way as to rule out all subjectivist and emotivist analyses of judgment. *De gustibus est disputandum* (ibid.).

¹²⁰ Here MacIntyre provides the notorious chess example: “Consider the example of a highly intelligent seven-year-old child whom I wish to teach to play chess, although the child has no particular desire to learn the game. The child does however have a very strong desire for candy and little chance of obtaining it. I therefore tell the child that if the child will play chess with me once a week I will give the child 50 cents worth of candy; moreover I tell the child that I will always play in such a way that it will be difficult, but not impossible, for the child to win and that, if the child wins, the child will receive an extra 50 cents worth of candy. Thus motivated the child plays and plays to win. Notice however that, so long as it is the candy alone which provides the child with a good reason for playing chess, the child has no reason not to cheat and every reason to cheat, provided he or she can do so successfully. But, so we may hope, there will come a time when the child will find in those goods specific to chess, in the achievement of a certain highly particular kind of analytical skill, strategic imagination and competitive intensity, a new set of reasons, reasons now not just for winning on a particular occasion, but for trying to excel in whatever way the game of chess demands. Now if the child cheats, he or she will be defeating not me, but himself or herself. There are thus two kinds of good possibly to be gained by playing chess. On the one hand there are those goods externally and contingently attached to chess-playing and to other practices by the accidents of social circumstance- in the case of the imaginary child candy, in the case of real adults such goods as prestige, status and money. There are always alternative ways for achieving such goods, and their achievement is never to be had *only* by engaging in some particular kind of practice. On the other hand there are the goods internal to the practice of chess which cannot be had in any way but by playing chess or some other game of that specific kind” (1985: 188).

This is the underlying objectivity that MacIntyre was looking for. Even if the fact that different moral traditions have generated accounts of virtues that significantly differ from one another, we should not conclude that, then, the cultivation of virtues is a subjectivist enterprise. The fact that each tradition presents its standards makes it so that there are objectively excellent ways to bring those about.¹²¹ Not only that, but different rules of excellence are likely to recognize the same valorization of virtues: for example, Bantu parents tell their children to hide the truth from strangers in fear of witchcraft; we are told by our (I'd say now grand-) parents to pretend we admire our old relatives' outfits. In both cases there is "an acknowledgment of the virtue of truthfulness" (193). Thus, MacIntyre concludes, "[p]ractices [...] might flourish in societies with very different codes; what they could not do is flourish in societies in which the virtues were not valued, although institutions and technical skills serving unified purposes might well continue to flourish" (ibid.). Virtues must already be valued and, thus, virtue cultivation must already be a practice in the first place, in order for moral paradigms to be virtue ethicist in this objective manner. What is crucial for our purposes is to note the relation between practices of virtue and their inherently historical dimension:

to enter into a practice is to enter into a relationship not only with its contemporary practitioners, but also with those who have preceded us in the practice, particularly those whose achievements extended the reach of the practice to its present point. It is thus the achievement, and *a fortiori* the authority, of a tradition which I then confront and from which I have to learn (194).

If we apply this specifically to virtue practices, we can see that it is an objective element of virtue practices that by participating in them one also enter into a relationship with

¹²¹ This also seems to imply tradition relativism. For an overview, see Kuna 2005.

their history, since it is from their history that one learns what it is that amounts to virtuous excellence, and what it does not.

It is important to note that MacIntyre does not wish to reduce morality to conventions,¹²² or mere traditional rule-following.¹²³ As stressed by Knight (1996), the idea is rather that since virtues are central to morality, then people should cultivate virtues, and they learn to do so by entering practices, “as practitioners come to find outside of themselves things that may be valued for their own sake and to devote themselves to pursuit of those goods” (Knight 1996: 21). Moreover, practicing in this MacIntyrean sense involves *reasoning*, not mere rule abiding. It is by engaging with practices’ particular form of reasoning that practitioners learn to pursue those internal goods that are proper only of those practices. Thus, argues Knight, “[t]his idea of practical rationality differentiates MacIntyre’s concept of practice from both conservative conceptions, of practices as subrational, and Foucauldian conceptions, of practices as discourses of power” (ibid.). In short, practices are those that generate a “commonly intelligible morality that facilitates the cultivation of virtue” (Knight 1996: 24).¹²⁴

¹²² See MacIntyre on “antiquarian conservatism” (1985: 222) as a vice that concerns possessing the adequate sense of traditions.

¹²³ on the Icelandic Sagas as a oral tradition, see Kristjánsson 1998.

¹²⁴ Knight inserts MacIntyre in the Aristotelian tradition, which “is that which justifies the life of practice and characterizes itself in terms of higher-level reflection upon the reasoning generated by traditional practices. This contrasts with both Enlightenment philosophy, which attempted to transcend practice and tradition, and Nietzscheanism, which shared the Enlightenment aim of liberating people from practice and tradition but denied epistemological certainty to any means of so doing. Aristotelianism is also, then, the philosophical tradition that maintains the rational validity of tradition and practice against other traditions that attempt to deny it” (1996: 24).

Even if it was not MacIntyre's intention to tie virtues to traditional conventions, there is still a sense in which virtues systematically arise out of a specific tradition, and in order for one to develop virtues in a progressive manner – that is, in a manner that is discontinuous with the moral status quo – the reasoning activity itself needs to be progressive rather than conservative, or backward-looking. It is not so clear how one can do that in MacIntyre's account, since reasoning about traditional virtues may just amount to revising traditional virtues in accordance with the conservative methodology that Ohlhorst has stressed (1.3). Is there an alternative? When it comes to identifying the virtues, MacIntyre's and Foot's approaches are two possibilities that are generally endorsed by neo-Aristotelians. That of MacIntyre seems inadequate to account for an understanding of the old virtues that is not a mere "reform" but a radical reconfiguration of both the list of virtues and of how to make sense of them. Let us see if Foot's naturalism offers a more progressive way to make sense of virtues.

2.2 Foot's Naturalism

According to Foot, moral judgments rest on rational considerations that amount to reasons to act morally, which is itself enough to motivate action (externalism). These rational considerations are natural facts about our belonging to the human species (e.g. the fact that we are social animals and that cooperation and solidarity are functional to survival, both individual and collective). Moral evaluation is a particular kind of a more general kind of evaluation, which is that of living beings and their function in relation to their species (Foot 2001).

That of Foot is an example of Neo-Aristotelian naturalism;¹²⁵ just as there is a particular type of flourishing for living beings (plants and animals) that is species-specific, there is a particular type of flourishing for human beings that is species-specific; that is, that has to do with the development of particular characteristics that are proper of the human species, such as the exercise of rationality (which coincides with the specifically human *ergon*). In short, Foot advocates for a moral theory that sees moral goodness as natural goodness, and moral vice as a natural defect.¹²⁶ That is, what makes a good person a good person is derivable from the kind of thing a human person is.¹²⁷

¹²⁵ There are different ways to interpret Aristotle's naturalism depending on the conception of human nature that he seems to employ. For example, one could read Aristotle through an essentialist interpretation (NE I, 7), where what is good for men is derived from their activities or specific function that consist in the exercise of reason (*logos*), which is specifically human, or through a biologicistic interpretation (NE VII, 6), according to which our first nature is the starting point of morality, that is, it needs to be guided by reason in order to develop a virtuous second nature.

¹²⁶ "Nobody would, I think, take it as other than a plain matter of fact that there is something wrong with the hearing of a gull that cannot distinguish the cry of its own chick, as with the sight of an owl that cannot see in the dark. Similarly, it is obvious that there are objective, factual evaluations of such things as human sight, hearing, memory, and concentration, based on the life form of our own species" (Foot 2008: 17).

¹²⁷ See Viggiano e Galletti on the neutrality of second-nature. In their overview, they note that Anscombe-Geach-Foot call for a second nature that is "neutral", that is derived from brute facts; McDowell, on the other hand, argues that second nature cannot be ethically neutral. Ethics is a product of moral education and the second nature that allows the agent to step back from her impulses and look at her motivations critically cannot be a view from nowhere; it must be an activity that takes place within a practical *logos*; that is, within those specific forms that second nature takes and is developed out of (McDowell 1998). Thus, there cannot be a neutral reconstruction of the foundations of ethics in natural terms. "Flourishing" and "fundamental human needs" are evaluative terms, not completely neutral terms.

According to Foot's naturalism "a part or aspect of a living thing is *naturally good* when it conforms to natural-historical descriptions of its life-form, and it's *naturally defective* when it doesn't" (Moosavi 2002: 5), where "natural-historical" judgments concern what has a function in the life of a particular organism.¹²⁸ Take the following judgments:

(a) "the male peacock has a brightly-colored tail"

(b) "the blue tit has a blue patch on its head"

only (a) is a judgment about natural goodness and defect, since only (a) has to do with the peacock's specific function as a peacock, which is that of attracting mates through its brightly-colored tail. On the other hand, the tit's having a blue patch on its head does not refer to the tit's specific function qua tit. Similarly, neo-Aristotelian naturalism extends judgments on the goodness and deficiency of natural functions to human beings qua human beings; "[t]hus, according to neo-Aristotelianism, moral evaluation of human character and action is the same kind of flourishing-based evaluation as the natural evaluation of an oak's roots or a peacock's tail" (Moosavi 2002: 6; Foot 2008: 18).¹²⁹ This need not be understood in strictly biological terms: neo-Aristotelians naturalists do derive what virtue is from empirical facts about human nature, but Foot herself warns against interpreting her account in terms of biological adaptation, or evolutionary biology. We should, rather, understand what she means by "function" through "everyday uses" of the word, not as a technical term (Moosavi 2002: 8). This is pretty vague; just how should we understand teleological claims on natural facts, such

¹²⁸ "Natural-history sentences, which Thompson also calls 'Aristotelian categoricals', speak of the life cycle of individuals of a given species" Foot 2001: 20, though Foot integrates Thompson's account with her distinction between judgment (a) and judgment (b) in what follows in the text.

¹²⁹ Neo-Aristotelians are not necessarily naturalists: see Dancy 2006, Andreou 2006, Millgram 2009 and Woodcock 2006.

as “the male peacock has a brightly-colored tail and this is good because it has to do with its function qua male peacock”? Foot writes that “‘natural’ goodness [...] is intrinsic or ‘autonomous’ goodness in that it depends directly on the relation of an individual to the ‘life form’ of its species” (Foot 2001: 18); that is, it is by looking at the life form of a living being that we can know what it is good for such living being, given its form:

We start from the fact that it is the particular life form of a species of plant or animal that determines how an individual plant or animal should be: the Aristotelian categoricals give the ‘how’ of what happens in the life cycle of that species. And all the truths about what this or that characteristic does, what its purpose or point is, and in suitable cases its function, must be related to this life cycle. The way an individual *should be* is determined by what is needed for development, self-maintenance, and reproduction: in most species involving defense, and in some the rearing of the young (Foot 2001: 22-23).

The fact that the male peacock has a bright-colored tail matters for the life of the male peacock, while the fact that the tit has a blue patch on its head is a non-teleological description, or a “statistical” one (Foot 2001: 23). This is because the absence of a bright-colored tail in the male peacock would be a defect, while the absence of a blue patch on the head of the tit would not.

As succinctly put by Moosavi 2002, we can summarize neo-Aristotelian argument for naturalism as follows:

(P1) Non-human life-forms are natural.

(P2) Evaluations of natural goodness in non-human life-forms are grounded in the nature of these life-forms.

(P3) Therefore, evaluations of natural goodness in non-human life-forms are natural.

(P4) Evaluations of natural goodness in human beings are of the same kind as evaluations of natural goodness in non-human life-forms.

(P5) Therefore, evaluations of natural goodness in the human life-form are also natural.

(P6) Evaluations of moral goodness are instances of evaluations of natural goodness in the human life-form.

(C) Therefore, evaluations of moral goodness are natural. (2002: 19-20)

How do we carve virtues out of this picture? As stressed by Hacker-Wright (2021), Foot's appeal to human nature has a twofold aim. On the one hand, it serves to tell us that what makes us human – our function as humans – is that we are rational beings; on the other hand, it serves to define what it is to fulfill our function; that is, to reason well. Practical rationality is species-relative, and our reasoning cannot do without what it is that we need as human beings, while also claim to exhibit practical rationality. This is important: according to Foot, nature is normative over our reasoning, since reasoning is our natural function as human beings. At the same time, though, this specifically human function allows human beings to be critical about our conceptions and interpretations of our nature. We need a conception of what a flourishing life is, in order to understand which virtues we need to cherish in order to flourish; but we can be wrong about what a flourishing life amounts to. Showing deficiency when it comes to making sense of our own function as human beings is a specifically human possibility.

If what virtues are derives from natural facts that prescribe what human flourishing is and how it is achieved, and if natural facts about what human flourishing is and how it is achieved are themselves subject to change, then we might have found a potential progressive element of VE. In this case, flourishing would be a dynamic source of normativity; we could make moral claims that are objectively true or false (e.g. “genocide is wrong” as objectively true) in virtue of what is good for human flourishing,

without a fixed account of what is ultimately good for human's flourishing. This potential openness would not just be revisionary, but potentially revolutionary, or status quo disruptive. However, this move is not so obvious. To begin with, we need to see if it is in principle possible for natural facts to be subject to change. If we understand them in evolutionary terms, then of course facts about nature are subject to *evolutionary* change. However, we have seen that a strictly evolutionary, or empirical reading of naturalism is not what Foot and other neo-Aristotelians have advocated for. So we need to see if we can make sense of natural facts in both a flexible and not strictly empirical manner.

In this respect, we have seen that naturalism implies moral realism, since it grounds moral facts on natural facts, thereby stipulating the existence of moral facts in the first place. This seems to suggest that there is an underlying moral ontology where there are such things as moral facts, and we can know them exactly as we come to know about natural facts. If this is the case, it is hard to see how entities in the world can have a flexible ontology, whereby facts about human nature can be themselves subject to change. Either we find a metaphysics that is flexible in this sense, or we need to do away with metaphysics. One noteworthy attempt to ground naturalistic claims in metaphysics is that of Swanton (2021): if her solution is able to both support naturalism – in her case, of second nature – and maintain nature's normative flexibility, then it would indeed be a good candidate when it comes to making sense of virtues in a progressive way. However, we shall see that it is not the case.

2.3 Naturalism of second nature and metaphysics: Swanton's alternative

A claim can be normatively justified (or unjustified), in which case it is an ethical claim. This need not imply that ethical claims require a supporting metaphysics in order to be

fully justified. In other words, the claim that “genocide is wrong” is a normative judgment, since it prescribes that genocide is to be avoided, and one can be an objectivist about it without needing a realist metaphysics to support such normative judgment (Fine 2002; Rosen 2017).

Ethical naturalism without metaethical naturalism would imply what Lott has called “the autonomy of ethics” (2020: 388); that is, that ethical claims are entirely concerned with what is normatively the case, and not with what is empirically the case. Lott notes that there are two ways to understand metaethical naturalism: one is epistemological, according to which we access ethical truth in the same way we access empirical truths (P4 of the argument for naturalism; Sturgeon 2002; Fitzpatrick 2008), the other is metaphysical, according to which ethical facts exist just like empirical facts exist in the world (Maguire 2015). Stronger naturalist versions of these claims would imply that there is a necessary connection between the ethical and the empirical. Lott notes that metaphysical naturalism is easily dismissed, as well as widely contested among naturalists themselves. On the other hand, the epistemological reading is widely accepted by naturalists, and it is thus harder to detach from a purely normative understanding of naturalism. Foot’s naturalism seems to rely on a strong version of the epistemic reading, according to which we access goodness through what we know of human nature. Now, on the one hand Foot warns us against reading “human nature” in purely empirical terms; on the other hand, if we do not understand “human nature” in purely empirical terms, we need to stipulate what humans are as entities that are not reducible to purely biological organisms. In short, we need to stipulate what kind of entity “being human” is; but in order to do that, we need an ontological apparatus that would make sense of humans as beings that are not merely biological organisms. In short, we would need a metaphysics in support of claims about natural goodness; that is,

we would need metaethical naturalism in its strictly metaphysical reading in order to ground ethical claims in natural claims.

This latter move is what Swanton has advanced in her book *Target-centered Virtue Ethics* (2021), where she supports McDowell's second-nature naturalism through Heideggerian ontology, grounding ethical claims about what is virtuous in the metaphysics of "Dasein," or "Being-in-the-world." She notes that

A major thrust of McDowell's work is to argue against restrictive conceptions of naturalism that tie the concept to the understandings of the natural (non-human) sciences. He does this primarily by claiming that the domain of ethics has what he calls a distinctive logos, one that is not reducible to the 'impersonal stance' of science. For him this distinctive logos is virtue ethical, marked by familiarity with the reason-giving force of the thick virtue and vice concepts (2021: 17).

Then, Swanton adds that just how one should understand this virtue ethical logos, as well as how it combines with the reason-giving force of thick ethical concepts is not clear. Her solution is to understand McDowell's second-nature naturalism as supported by hermeneutic ontology, so as to give proper metaphysical ground to thick ethical concepts, finally explaining in what sense they are reason-giving and action-guiding and ultimately generate virtue ethics' particular logos.

Swanton (2021) characterizes the virtue ethical family through the following two theses:

(A) Thick Concept Centralism (TCC), according to which "claims that the general concepts of right and ought are not logically prior to and independent of reason giving thick concepts that favour or disfavour responses (such as action)" (5)

(B) The Centrality of Virtuousness (CV), according to which “rightness and/or excellence are understood through the virtue/ vice notions” (ibid.).

Where “TCC does not entail CV for one basic reason: though the thick concepts provide evaluative shape to non-evaluative reality and thus demarcate an evaluative reality, they do not in themselves connote excellence or even pro tanto rightness.” (Swanton 2021: 5). Moreover, TCC combined with CV (TCC/CV) does not make virtue concepts alone central; rather, “what is central are the thick concepts in general combined with notions of rightness and/or excellence” (ibid.), e.g. being generous, or doing something generous, combine with notions that see being generous, or doing generous acts as showing rightness and/or excellence. Finally, TCC/CV does not itself require that VE offer an account of right action.¹³⁰ It is important to note that Target-centered VE adds a third characterization:

(C) TC, Target Centeredness, “according to TC what counts as good and correct forms of responsiveness to the world, whether to things, persons, animals, situations, whether in relation to feeling or action, is determined by the targets of the virtues, relative to kinds of responses, whether they be respect, love, creativity, promotion of good, and so on” (7).

Now, we saw that for naturalism to offer a solid basis for ethics it is hard not to understand it through natural sciences, yet doing so raises a number of problems. McDowell has famously argued against accounts of naturalism that restrict the understanding of natural facts to the domain of natural sciences. Instead, he argues that the domain of ethics is a distinctive logos, not a sort of view from nowhere typical of science, and this distinctive logos is virtue ethical to the extent to which it is the domain of responses to reasons to act in accordance with the requirements of thick virtue

¹³⁰ As a matter of fact, she is skeptical of the notion of right action: see Swanton 2001.

concepts. In Swanton's terms, McDowell's effort is that of combining suitable versions of TCC with realism and naturalism (Swanton 2021: 17), which is better understood by drawing insights from Heidegger's hermeneutic ontology.

The first thing to consider is that, for Heidegger, "[w]e are essentially practical beings rather than disengaged theoretical observers of the world" (2021: 18). Taylor considers the perspective of the view from nowhere and notes that "Heidegger had to struggle against this picture to recover an understanding of the agent as *engaged*, as embedded in a culture, a form of life, a "world" of involvements... Understanding ontology then requires on this view an understanding of the nature of Dasein—the human way of being in the world." (in Swanton 2021: 19, my italics).¹³¹ In light of this, Heidegger's hermeneutic turn consists in acknowledging that "truth and knowledge presuppose a context of meaningfulness through which we understand the nature of entities" (18), where "meaning" for Heidegger involves the intelligibility of an entity as an entity in a web of relations, which are thus called "relations of significance."¹³² In Swanton's words, "human beings are described by Heidegger as Dasein—literally being there—that is to say, having a being which is relational; being in a world with others and with which we are engaged. Hence, to understand the nature of entities we need to understand the manner in which human beings engage with the world" (18).

Swanton characterizes Heidegger's hermeneutical ontology through 7 theses (2021: 19-20):

¹³¹ As stressed by Swanton, "The way human beings dwell in the world is one of absorption and engagement in a social context. It is not primarily one of detached spectator, which in its extreme form is a pathology rather than a recipe for maximal objectivity" (Swanton 2021: 24).

¹³² See also Taylor on "engaged agency" (2006).

Thesis 1. For an entity to exist as something we can talk about it must have a certain mode of existence; it must have “being.” For example, it exists as a piece of equipment, or as an artwork, or as an object of scientific observation. In Heidegger’s words, “Being is that in virtue of which entities are entities,” or a “mode of existence”. Being is not a “class or genus of entities” (B&T sect. 38).

Thesis 2. For an entity to have being one must be able to intend it: we must have intentional access to that entity. To intend an entity is for some intentional state to be about or directed at something, and thereby we make sense of them. Hence there is two-way ontological dependence. “Being is entities making sense (to us) as entities.”¹³³ This implies that what entities are depends on our making sense of them “as people *engaged* in the world in various ways” (Swanton 2021: 19, my italics).

Thesis 3. Intentional access to entities (intending entities) is gained through frameworks of significance. Such frameworks of significance are constituted by what Heidegger calls logos; logoi such as the natural sciences and social sciences, biology, theology, carpentry, art, ethics. It is through logos the phenomenon is ‘put together’ with logos.¹³⁴ In this way an entity is ‘disclosed’. This also implies that “[t]here is no independent access to what Heidegger calls the ‘factum brutum’, by a ‘pure beholding’” (ibid.).

Thesis 4. When entities are ‘put together’ in this way they constitute what Heidegger calls a ‘worldhood’ such as the worldhood of equipment; the worldhood of ethics. A worldhood for Heidegger is ‘that referential totality which constitutes significance.’¹³⁵

¹³³ ibid.

¹³⁴ Being and Time, sect. 28, 50.

¹³⁵ Being and Time, 160. See also 114–22.

“Since logos is the medium through which phenomena are made manifest, the ontology of entities (their worldhood) cannot be conceived as independent of our modes of intentional access, and in particular frameworks which make entities significant to us” (Swanton 2021: 20).

Thesis 5. The disclosure of entities in a worldhood enables things to matter to us; makes them available for us to engage with them. Unconcealment¹³⁶ of entities through logos makes them ‘available for comportment’ or involvement. Worldhood then is a worldhood of involvements.

Thesis 6. There are several frameworks (logoi) disclosing entities that are not reducible to the scientific or to any other privileged framework (such as the worldhood of equipment). Thus, entities have a variety of modes of existence—a variety of modes of being, and there are thus several worldhoods. We are engaged with the world, things matter to us, in a variety of ways. We should not privilege one way of mattering, observing the world in the manner of scientific observation.

Thesis 7. The disclosure’ or ‘unconcealment’ of entities through frameworks of significance precedes epistemology. (2021: 19-20)

That is, an entity exists in the world, in its relations of meaning, and knowing it, or understand it takes different forms. As explained by Heidegger, “[w]e therefore distinguish not simply terminologically but for reasons of intrinsic content between the discovery of an entity [Entdecktheit eines Seinden] and the Disclosure of its being

¹³⁶ According to Wrathall (2011) ‘unconcealment’ (the usual translation of Unverborgenheit) is Heidegger’s preferred translation of aletheia (openness or truth in his special sense).

[Entschlossenheit seines Seins]. An entity can be discovered whether by way of perception or some other mode of access, only if the being of this entity is already disclosed”¹³⁷ (in Swanton 2021: 20).

Swanton then draws a parallel between Heidegger’s hermeneutic ontology and McDowell’s claim that there is a distinctively ethical mode of access to the real. How is this ethical mode of access to be understood? “Properties such as being virtuous are attributed to entities within a framework of significance relations, such as that of ethics. Ontology then is not simply the entities that exist, but entities that exist *as meaningfully engaged with* by human beings. In this way openness through the framework of significance relations constitutes the “being” or ontology of an entity. As McDowell suggests there may be more than one mode of openness and thereby more than one mode of being, modes such as the mathematical, the ethical, and so on” (Swanton 2021: 21, *my italics*).

In short, what is common to both Heidegger and McDowell is that “[w]hat is crucial is that rather than things or facts being “brute” (a term used by both McDowell and Heidegger) they are intentionally accessed within a meaningful context” (22).¹³⁸

¹³⁷ Cited in Sacha Golob 2014: 87.

¹³⁸ Epistemic access as distinguished from more “basic” kinds of access. “‘Understanding’ like ‘access’ is ambiguous between intelligibility and justified belief; between our intentional access to properties of a certain broad type such as the ethical, which is needed for the ethical to exist, and an epistemic notion where our openness to, for example, virtue or virtuous action has to be justified as accurate. As McDowell puts it, once our eyes are opened to the practical ethical world (constituting the intelligibility of ethics as something which is open to us) epistemology takes over: ‘Thereafter our appreciation of its detailed layout is indefinitely subject to refinement, in reflective scrutiny of our ethical thinking’ (22).

The move that Swanton is suggesting is the following: in order to understand McDowell's attempt to dismiss biologicistic interpretations of naturalism, while maintaining the importance of what kind of beings humans are when it comes to make moral claims, we need to move away from a conception of human beings qua human beings, towards a conception of human beings qua Dasein. What provides meaning to actions is a virtue ethical logos, where relations of significance take the form of thick ethical concepts. What provides the action-guidance of seeing a person in need in favor of helping her is the fact that this person and us both exist in a network of significance whereby helping her makes ethical sense because it makes metaphysical sense: she exists as a person-in-need and I exist as-a-person-who-can-help. Swanton also notes that all this resonates well with MacIntyre's idea of practices and traditions:

“[a]nother basic feature necessary for logos is being situated in a culture within a tradition [...]. McDowell's position [...], is that immersion in a tradition is essential for intentional access to entities through the various modes of access (logoi). Logos in general, whether that of ethics, science, equipmentality, is immersed in tradition.” (2021: 27).

To elaborate on this – in place of McDowell – she draws a parallel with MacIntyre's *After Virtue*:

“According to [MacIntyre] we necessarily act within ‘settings’ [...]. It is essential to the notion of a setting that it has a history. To understand an action ‘we have to understand its history.’ On a large scale the setting is our culture, and the history of our culture is our tradition. More specifically, intentional access to the normative properties

of actions through the logos of ethics, most abstractly specified by the ‘space of reasons’ , essentially involves settings, social and historical location, and thereby tradition. As MacIntyre puts it, ‘ . . . all reasoning takes place within the context of some traditional mode of thought.’ (ibid.)

Thus, if we ground McDowell’s second-nature naturalism in hermeneutic ontology, we avoid the risk of supporting ethical claims through biologicistic interpretations; but we are not in the position to avoid the inherent inflexible nature of virtue ethics’ logos, whereby virtues’ action-guidance is understood through a network of relations of meaning that make sense *exclusively* when embedded in a particular tradition of meaning. In other words, one can act virtuously in the context of a moral paradigm only because one is embedded in that moral paradigm. But there are no internal tools to be even in principle in the position to transcend, or disrupt that moral paradigm. Even bleaker does the picture become when one grounds all this in metaphysics: it is just how entities exist, that is, as embedded in a moral paradigm. Virtues exist as relations of virtuousness between us and the object (or, in Swanton’s terms, the target) of virtues. But all that is already determined by the moral paradigm itself, and just how one could exist as any other type of entity (e.g., as having no meaningful relation with the moral paradigm), is hard to conceive.

We need to dismiss metaphysics as something that provides explanation as per what it is that guides actions in virtue ethics. It is true that there is an inherent tension in VE that on the one hand prescribes objective ways to act virtuously, and on the other places huge importance on contexts and traditions. This is why VE is particularly inhospitable to radical moral change. However, we need not ground the virtuous thing to do in ontology. I will conclude this section with a possible way to argue for the possibility of

making progressive ethical claims without the necessity of a supporting metaphysics. In the last section I will expand on what VE would need in order to make room for such possibility.

2.4 Ethical Naturalism without Metaethical Naturalism

We have seen that Lott does not consider Foot's implicit endorsement of the metaphysical reading of naturalism, and argues that Foot's metaethical commitment is both epistemological and incompatible with the autonomy of ethics, understood as the independency of ethical claims from empirical claims. He notes that

even if we grant Foot's core claim that the grammar of goodness applies to human being and that moral goodness is natural goodness in the human will, there remains a question: Is the way in which we come to understand human good(ness) the same way that we come to grasp the good(ness) of plants and non-human animals? If Yes, then it seems we must reject the autonomy of ethics. However, it is possible for an Aristotelian naturalist to answer No, in which case ethical thought might still be autonomous. The basic idea would be that the grammar of goodness applies to human beings, but we cannot grasp what is naturally good in the will of human beings if we restrict ourselves to the picture of "the human" available via the empirical sciences (2020: 395).

As noticed before, if access to ethical knowledge is not entirely attributable to access to empirical knowledge, then we need a picture of "human beings" that is not purely reducible to empirical facts about human organisms. This, however, would need a metaphysics of what humans are as entities. Lott argues that we can make sense of

Aristotelian naturalism without the need for the epistemic reading of metaethics; but as I have tried to show, doing away with a purely empirical reading of “human nature” or “human beings” while also maintaining the *objective* element of naturalism, would require the metaphysical reading of naturalism. The question then is the following: is there a way to make sense of “human beings” and of “human nature” in purely normative terms, without the need for a metaphysics of “human beings” and of “human nature” that supports these normative terms, while also preserving elements of objectivity for which we could legitimately claim that, say, genocide is wrong, whether there is a supporting metaphysics or not? The answer is yes, and it is an important answer for the purpose of this research for the following reason: VE is inherently conservative to the extent to which it defines what virtues are starting from either tradition or nature, or both. The presence of moral traditions is not itself a conservative element: what is conservative is the relation between virtues and moral traditions, that sees virtues as necessarily arising out of traditions and not, say, in opposition to traditions; that is, in a *discontinuous* manner. Similarly, conservative claims about virtues have generally been supported by natural claims; i.e., claims that reinforce the moral status quo of traditions rather than oppose it.

Now, the underlying intuition in ethical naturalism is that there is something about being humans that makes it so that some things are good and others are bad. It is this very intuition that drives progress, both at the structural and at the individual level: the recognition that there is something good that we are not pursuing, or facilitating; something bad that we are, on the other hand, actively pursuing, or implicitly facilitating. Thus, it is an important intuition to maintain; can we do so without affirming naturalism’s problematic implications? Wilson (2010) has argued that we do not need moral realism to affirm the truth of progressive moral claims such as “a society without slavery is better than a society with slavery,” since we can frame change for the

better in terms of narratives that have proven to be true: “[t]ruth ... is opinion that is the outcome of certain definite procedures and that has survived certain definite tests” (2010: 113). Just like scientists test their paradigms, we can test our moral paradigms – which she understands as our expressions of positive or negative attitudes towards some moral matter – to see if they are “consistent with new and old data” (ibid.), and then decide to keep them or change them.¹³⁹ Here the objective element is provided by scientific-like tests on our moral claims and, given our discussion so far, we can see how this would be an instance of epistemic naturalism, whereby we can access moral truths just like we access empirical truths. Moreover, the narrative element is what prevents the need for moral realism. Now, if we add the stronger naturalistic claim according to which it is human nature that is normative, we can see that what provides normativity are narratives about human nature. This would prevent purely biologicistic readings of human nature and, thus, problematic normative frameworks. However, it is far from clear whether narratives can truly present the objective element that we are looking for. That of Wilson is a description of how certain instances of radical moral change – mostly societal – can take place, which is consistent with our discussion of MPS through the analogy with Kuhn’s paradigm shifts in science: but there must be something that makes a scientific paradigm more accurate than another. Similarly, there must be something that makes a moral paradigm more accurate than another. There must be some *reality* that is accessed when the paradigm is accurate. Thus, it is not so clear whether Wilson’s account truly dismisses moral realism, if not by affirming, with expressivists, that moral claims are just the expression of one’s moral attitudes, and maybe some moral attitudes are tested as more useful than others.

¹³⁹ Wilson argues that in this way we can accept what is right about expressivism, while dismissing its most problematic implication that sees no moral objectivity in support of moral claims.

More recently, Severini (2021) has argued that a promising way to argue for the possibility of affirming the truth of progressive moral claims without the need to affirm the existence of moral facts to support them is to frame the debate in terms of moral understanding. There are no moral facts; but what we can do as individuals is to tune into reality and transform our moral outlook “into some other moral outlook that we deem to be better” (2021: 102). To put it differently, there are no moral facts, but there is a *normative reality*.¹⁴⁰ Severini employs Murdoch’s M&D example (Chapter I, section 2 of this dissertation) to show that M’s enhanced moral understanding of D is an instance of non-factual moral progress: “[t]he activity in which M has been involved is that of refining the way in which she sees the world: she perceives previously unnoticed features of D, she gains a novel appreciation of how various elements are related to each other and, thus, she gets a better grasping of the situation” (ibid.). The objective element here is the fact that if one attunes to “unnoticed features” of the situation, then one learns to look at the world through different lenses. What distinguishes “better” lenses from worse ones is “whether we can use our capacity for moral reflection to produce answers to certain moral demands or problems that we can explain and justify to others” (2021: 103). It is in this sense that change for the better is not an arbitrary process: one can morally evaluate it on the basis of reasons for action, reasons for inaction, duties, moral demands, etc. This might sound just like another way to claim that we form narratives that seem coherent with available moral data, as Wilson argued. The difference between narratives and moral outlooks is that in narratives the objective element is still provided by access to facts – e.g. moral and scientific data – while the objective element in outlooks is provided by attunement to normative reality. Silvia Caprioglio Panizza (2022) explains normative reality in terms of “affordances”: if you

¹⁴⁰ Recall that naturalism does not endorse the fact/value distinction.

not only observe a sick person, but *attune* to a sick person, then you will see her through the normative affordance “help-able.” That is, the reality of that person¹⁴¹ prescribes what one ought to do. The objectivity here is the fact that normative reality is not disclosed unless one attunes to it free from one’s own self-concerns and biases (unselfing). If I see the affordance of a sick person as “rob-bable” given she is not able to react to external attacks, I have not attuned to her, but only to my own self-concerns (e.g. “I need cash”). Similarly, what provides normativity in Severini’s understanding-based account is something that is not within one’s self. In this sense, it is less subjectivist than Wilson’s expressivism. If we add the naturalistic claim that what provides normativity has to do with what human beings are, or what human nature is, we can see that in order to understand in the first place what human beings are, or what human nature is, we need to look at what is *not* ourselves. The human phenomenon in its complexities, in its otherness. For there to be something objectively good or bad one need not see human beings as metaphysical realities, but as normative realities. In the next and final part I will provide an example of what this might mean.

3. A Progressive Moral Outlook

To recap, we started from the following observation: VE’s commitment to the *development* of virtuous character traits makes VE a particularly dynamic theory, since it makes it a theory that relies on constant moral change towards achieving virtue. On the other hand, VE accounts hardly ever mention MPS as a particularly noteworthy dynamic of moral change. I argued that this is likely due to the fact that virtue ethical

¹⁴¹ Or just “an other”.

theories focus on moral development, which is status quo reinforcing, while do not seem to account for the possibility to transcend ordinary moral development, that is, disrupt the moral status quo. In particular, in neo-Aristotelian moral development agents develop in accordance with an existing moral paradigm and develop into fully virtuous agents in accordance with that moral paradigm. It is, however, important to be able to account for possible changes in moral paradigm; specifically, it is crucial for moral education to elicit the capacity for potential disruptions of paradigms. In relation to this latter point, we have seen that, as stressed by Rehen and Sauer (2024), moral education is much more likely to impede moral progress rather than to cause it. Thus, neo-Aristotelian VE need to rely on other devices in order not to impede moral progress. As noted by Ohlhorst, VE methodology is inherently conservative, since it focuses on what we already think rather than on what we should be thinking independently on how we have understood it so far. Therefore, it seems highly implausible for VE to be progressive, unless it radically changes in methodology.

We have then considered two main trends in VE theories that conceive of virtues as tied to traditions (MacIntyre) and as tied to facts about human nature (Foot), noticing that both trends seem to be particularly inhospitable for the possibility of moral status quo disruption. I then argued that an ethical naturalism without metaethical naturalism would allow for virtues to be conceived in terms of what is good for human beings qua human beings, without necessarily tying conceptions of what is good for human beings to inflexible understandings of what is natural for human beings. In this way, one need not de-naturalize neo-Aristotelian VE theories from their core assumption, that sees virtue development as central, and an idea of flourishing as normative, because virtue cultivation can be both central and open-ended, and an idea of flourishing can itself be subject to paradigm shifts. However, these two latter points have yet to be defended. As a matter of fact, it is far from likely that neo-Aristotelian accounts would in principle

accept either of these caveats. I will end by expanding on the following points: 1, if VE theories, be them neo-Aristotelian or not, do not structure moral development as functional to the development of a progressive moral outlook, then they cannot be progressive; 2, VE theories, be them neo-Aristotelian or not, *should* structure moral development as functional to the development of a progressive moral outlook; and 3, it is possible for VE theories to do so.

3.1 Moral Development as Epistemic Development

It is time to defend Argument 2:

1. if VE shifts the focus on epistemic virtues, rather than moral virtues, then VE can be a flexible moral theory
2. if VE rests on a non-realist notion of “flourishing”, VE can provide a dynamic account of flourishing
3. if VE can in principle be flexible, then it should be so
4. VE can shift the focus on epistemic virtues (virtue epistemology)
5. VE can rest on a non-realist notion of “flourishing” (ethical naturalism without metaethical naturalism)

=> VE can in principle be flexible

=> VE should be flexible

In 1.1 we explored neo-Aristotelian moral development as a way for individuals to develop into morally mature agents. What I wish to stress now is the fact that current accounts of neo-Aristotelian moral development significantly rely on the development

of epistemic character, but do not explicitly argue for it, nor expand on the relevance of developing epistemic virtues *in order to* develop moral virtues.

In particular, we saw that neo-Aristotelian stages are: moral indifference, akrasia, enkrateia and proper virtue. We also saw, although not explicitly, that in order to go from one stage to the next, some sort of epistemic goal needs to be achieved. For example, in moral indifference people are not able to determine what the virtuous thing to do is and why (Sanderse 2015: 7), but in akrasia, the second stage, they are. Akratic people “see the attractiveness of a life of virtue, yet they lack the ability to control their passions” (1.1) that still go in non-virtuous directions. The moral outlook of moral indifference is significantly different from that of akrasia: in the former, agents do not see virtuous requirements, while in the latter they do. The epistemic element of the moral outlook is the one that allows agents to see the attractiveness of the virtuous thing to do; the attractiveness of “moral reality,” in Murdochian terms. Different is the case for the moral outlook of the akrates and that of the enkrates: they are pretty similar in their being able to frame situations in a way that allows them to see the virtuous requirement, and they are also pretty similar in the way they emotionally relate to the virtuous requirement, that is, as something they would rather not pursue. Yet the enkrates has deeper moral understanding compared to the akrates; this is why she manages to pursue the virtuous requirement nonetheless. This deeper moral understanding is what ultimately distinguishes the moral outlook of the enkrates from that of the akrates, and it is an epistemic gain proper of the enkrates stage. Finally, the moral outlook of the fully virtuous is one that frames situations in accordance with her even more accurate moral understanding of situations, in a way that allows her to act on

the virtuous requirement and to wish to do so.¹⁴² In short, when it comes to moving from one stage to the other, there is always some epistemic achievement that directly concerns the moral content of one's moral outlook. In other words, what it is that one values and what it is that one finds action-guiding is dependent on how one makes sense of the moral dimension of life. It is in this sense that epistemic development is functional to moral development.

Now, if neo-Aristotelian moral development openly acknowledged this dependency of moral development on epistemic development, it would be in the position to form not only individuals that are perfectly capable of developing into morally mature agents, but also be in the position to form individuals that are in principle equipped with progressive moral outlooks; that is, that are at least in principle open to the possibility to transcend their own moral paradigms, and are equipped to re-adjust their moral outlooks according to new, unexpected moral paradigms. Standard, neo-Aristotelian moral development is liable to the charge of conservatism because it does not explicitly contemplate the possibility that agents develop epistemic characters that are progressive to the extent to which they are able to transcend existing moral paradigms. Sanderse (2015) and Annas (2011) did claim that the last stage, that of full virtue, is open-ended and, thus, one need not decide in advance what it is that amounts to full virtue but can discover it through her stage advancement. But this is not enough to construe a progressive account of moral development within a VE framework, because this solution does not contemplate the possibility for agents to develop epistemic characters that are in and of themselves capable of operating paradigm shifts at whatever stage they are. Going back to the ladder analogy, a progressive account is not one that claims

¹⁴² Which does not imply that the life of virtue is a happy life because the fully virtuous is perfectly fine, say, sacrificing her life if that is the virtuous thing to do; nor that, for the same reason, the life of virtue amount to miserable slavery to virtuous requirements, as I argue in Liberti 2022.

that the ladder of moral development is potentially infinite, but one that aims at equipping agents with the tools to *change* the ladder they are climbing, bringing about the *real* moral progress that Wittgenstein had in mind, in opposition to mere growth, or the scratching of a finger. If VE accounts and, in particular, neo-Aristotelian accounts of moral development can integrate their theoretical framework with the development of epistemic virtues that enable moral paradigm shifts, then VE accounts would finally be able to rely on a progressive methodology and, thus, have something substantial to say about progressive changes like MPS, as well as, importantly, be able to provide progressive action-guiding responses to pressing moral issues without turning back to moral traditions, what has been the case so far, how we have conceived of something in the past, and so on. If VE accounts manage to do that, then two important things would follow: (a) specific epistemic virtues would need to be included in accounts of moral development, and (b) accounts of “flourishing” and “human nature” would be dynamic, to the extent to which they would themselves be subject to moral paradigm shifts. In what follows I will argue that this is what VE accounts should do. In particular, I expand on both (a) and (b), and conclude that if this epistemic move is made, then VE accounts would be perfectly able to provide agents with progressive moral outlooks.

3.2 Epistemic virtues for a Progressive Moral Outlook.

To begin with, what are “epistemic” virtues? The current debate in virtue (and vice) epistemology revolves around two approaches, “reliabilist” and “responsibilist” (Baehr 2011; Croce 2020; Kidd 2021), which differ in what they take an epistemic virtue/vice to amount to. According to reliabilists, an epistemic virtue is a faculty (also “faculty-based”, Baehr 2011) just like any other cognitive faculty (Sosa 1980; Goldman 1998;

Greco 2000), and an epistemic vice expresses, thus, an *intellectual deficiency* or failure just like being bad at calculation or having low attention span. On the other hand, responsibilists understand epistemic virtue as an expression of the agent's character traits and/or mental attitudes (also “character-based”, Baehr 2011) and see epistemic vices as *vicious character traits* and/or *mental attitudes*. The label “responsibilism” comes from the further specification according to which an epistemic vice can be rightly called so if the agent is responsible for having or managing it (Zagzebski 1996; Hookway 2003). For this reason, Battaly (2015¹⁴³) labels these approaches “motivation-based”, which also nicely affiliates them with Aristotelian virtue ethics according to which it is a misconception of the good that motivates the vicious agent to act, (NE 1150b36-37; Tanesini 2021a; 2021b)¹⁴⁴. In what follows, we can think of epistemic virtues as epistemic habits that are virtuous to the extent to which they facilitate the understanding and the development of moral virtues. This definition is more akin to responsibilist accounts, since it implies that epistemic virtues and vices determine one's epistemic character. Here I would add that not only epistemic virtues and vices determine one's epistemic character, but one's epistemic character is what goes to construe one's moral outlook.

Now, “epistemic virtues” is a tricky name. We could just say “epistemic habits”, but by calling them “epistemic *virtues*” we end up conflating intellectual and moral virtues.¹⁴⁵ This is because the name “epistemic virtues” implies that there are some epistemic habits that are also *virtuous*; that is, morally “good,” and it is not clear how it can be so. It is perfectly possible that a highly intelligent person is morally corrupt; what then is

¹⁴³ See also Kidd et al 2020.

¹⁴⁴ Although this affiliation has been recently challenged, see Kotsonis (2021) for an argument that classifies Aristotle as an epistemic consequentialist rather than a motivation-based virtue epistemologist.

¹⁴⁵ Which, for some, is problematic. See Driver 2003.

the relation between epistemic and moral character? If we understand epistemic virtues as mental habits that are functional to moral virtues to the extent to which they provide moral understanding and, thus, construe one's moral outlook, then the relation amounts to just that: epistemic habits can be virtuous when they enable one's understanding and development of moral virtues. Epistemic virtues in this sense are not reduced to cognitive abilities, nor to knowledge that is merely literacy. In other words, one's epistemic character does not amount to how knowledgeable one is on, say, philosophy or mathematics, nor on how high is their IQ. Rather, it amounts to the person's mental attitudes towards knowledge (Baehr 2011). When it comes to the relation between epistemic and moral virtues, we can say that one's epistemic character determines one's attitude towards moral issues. There can be ways to frame situations that are more or less progressive, to the extent to which they are more or less open to MPS, and to the extent to which they are more or less able to re-adjust one's value system in virtue of a MPS. A progressive moral outlook, thus, requires an epistemic character that is in principle open to MPS: this epistemic character is also morally significant, since it is what allows agents to undergo, and flexibly adjust to MPS. This does not imply that there is a specific content to those epistemic and moral virtues that make for a progressive moral outlook. To put it differently, it is the *relation* between agents and new moral possibilities that make for a progressive or conservative moral outlook: progressive if open-minded and flexible about new moral possibilities, conservative if narrow-minded and inflexible about new moral possibilities. In this picture, whether agents are also, say, educated at an Ivy League University or have worked throughout their entire life without the material possibility to satisfy intellectual curiosities does not really matter. What makes for their moral outlook when it comes to the possibility of radical change is in what relation they stand towards moral innovations. This is why

epistemic virtues need to be cultivated. I will end with a proposal on which epistemic virtues we may wish to develop in order to construe progressive moral outlooks.

The first epistemic virtue that is likely to be required in order to develop a progressive moral outlook is phronesis. In NE Aristotle first considers the single virtues of character (NEI-V), then goes on to show how these single virtues of character cohere (NE VI): it is phronesis that allows the single virtues to form a unified character¹⁴⁶, by making them part of an integrated whole. In other words, one may be sometimes courageous, but to *know* when, how and to what extent it is right to be courageous in every situation is to possess phronesis.¹⁴⁷ Thus, phronesis in its original sense can be understood as the epistemic virtue that integrates moral virtues into the phronimos' character. Nowadays we can find different, competing models of phronesis,¹⁴⁸ that draw from this original Aristotelian intuition that sees phronesis as having a key role in the development of a virtuous character, whereby it is not enough for one to merely “collect” moral virtues. Something else is required. Viewed in this way, phronesis takes the form of a “meta-virtue”,¹⁴⁹ that is, of a virtue that is epistemic since, as it is its original form, it has to do with the epistemic character of a person, something akin to what we consider an

¹⁴⁶ on the unity of virtue in Aristotle, see Wolt 2022; for a current account of phronesis and the unity of virtue, see DeCaro, Vaccarezza and Niccoli 2018.

¹⁴⁷ this is why it is generally taken to be an epistemic virtue, but the issue is tricky. Going back to the conflation problem (Driver 2003), phronesis seems to be the virtue that does inevitably combine both epistemic and moral aspects.

¹⁴⁸ See for example the Jubilee model Darnell, Fowers, and Kristjánsson, 2022; Swartwood 2013 and Tsai 2022 on phronesis as a form of expertise. For a comprehensive model of phronesis as the only virtue (monism) understood as ethical expertise, see the ARETAI model, forthcoming.

¹⁴⁹ See in particular the ARETAI model (note 146), although not a strictly neo-Aristotelian.

“outlook”:¹⁵⁰ that is, it has to do with how one codifies information, how one interprets them, how one finds some more salient than others, how flexible one is concerning one’s understanding, and so on. In particular, phronesis is the meta-virtue that specifically deals with perceiving salient features of the situation and coordinating the single moral virtues to act on such requirements.

Now, imagine a phronimos undergoing a MPS. The Gestalt switch discloses to the phronimos new affordances; that is, new ways of making sense of and dealing with the situation at hand. In short, a normative reality that is different from the one she used to see is revealed. This Gestalt switch causes a cognitive dissonance, where the existing moral paradigm, along with the existing beliefs of what a flourishing life amounts to are shaken to the core. The ‘perceptual’ function of phronesis enters the picture, enabling a new, accurate construal of the situation. This new, accurate construal of the situation calls for readjustment of one’s moral outlook, which is exactly what happens in virtue of a MPS that has taken place. However, we need to stress two factors: first, it would be unrealistic to charge phronesis with the entire job of revising one’s moral paradigm and adjusting to a new, radically different one; second, it would be unrealistic to claim that one needs to be a full-blown phronimos in order to undergo MPS. For these reasons, further epistemic virtues are needed.

In an influential taxonomy Baehr (2011: 21, revised in 2021) groups epistemic virtues into six categories that consider how these virtues can help meeting challenges to moral inquiry. Among these groups, there are two that are particularly promising candidates when it comes to forming moral outlooks that are at least in principle transformative; that is, able to welcome MPS and transform one’s value paradigm accordingly.

¹⁵⁰ See Chapter II of this dissertation.

The first group considers “virtues of intellectual wholeness,” such as intellectual integrity, honesty, humility, transparency, self-awareness, and self-scrutiny, which all play a key role in questioning one’s beliefs in light of new evidence (2011: 19). The second set includes “virtues of mental flexibility”, such as imaginativeness, creativity, intellectual flexibility, open-mindedness, agility, and adaptability. These virtues are required when one “confronts a subject matter that [...] is foreign to her usual way of thinking. Here what is needed is an ability to ‘think outside the box’.” (Ibid.)

While virtues of intellectual wholeness enable self-scrutiny and can bring about a revision of one’s moral beliefs, virtues of mental flexibility are those that enact the disintegration of previous moral paradigms and allow to conceive of new, different ones. In particular, creativity, imaginativeness and open-mindedness allow an agent to “confront a subject matter that is in itself extremely complex and demanding or that is simply foreign to her usual way of thinking” (Rees & Webber: 20). Thus, they are crucial when it comes to facing the cognitive dissonance that is produced when new affordances are disclosed, and to find new construals.

Finally, it is also worth considering epistemic justice as a potentially transformative epistemic virtue. In Miranda Fricker’s (2007) notorious account, epistemic justice comes in two forms: testimonial and hermeneutical. While testimonial justice amounts to “reliably neutralize the impact of prejudice in [...] credibility judgements,” hermeneutical justice is supposed to correct unjust attitudes and behaviors whereby “someone has a significant area of their social experience obscured from understanding owing to prejudicial flaws in shared resources for social interpretation” (2007: 147). As a corrective virtue, it consists in considering the possibility that an interlocutor’s difficulties when it comes to articulating reasons and sharing conceptualizations are not due to epistemic shortcomings of some sort, but to a lack of access to collective hermeneutical resources. In relation to MPS, epistemic justice allows people to

overcome prejudices and marginalizing attitudes when re-shaping their moral outlook, as well as to develop higher levels of social awareness and grant due credibility to marginalized members of a community, while also considering their own status and conditions as epistemic agents – likely, privileged ones. Since epistemic justice considers individuals in their broader social and cultural context, epistemic justice as a transformative virtue has the potential to lead to instances of both individual and societal moral progress.

3.3 Dynamic flourishing

If one is equipped with a progressive moral outlook, then a dynamic account of flourishing follows, since flourishing itself would be subject to possible MPS. “Flourishing” is one of the normative elements of VE:¹⁵¹ one should be in the position to flourish;¹⁵² and one ought to X because X is the flourishing option. But what flourishing amounts to is a long-standing debate. The point I would like to make is the following: if one develops one’s epistemic character, then one need not stipulate what flourishing amounts to in advance. Flourishing may be discovered; it may be the epiphanic element of MPS, when they take the form of epiphanic moral conversions. And that itself may be subject to constant, more or less radical change. One may learn that, according to a particular moral paradigm, flourishing amounts to “becoming the best version of yourself”,¹⁵³ yet discover that becoming the best version of yourself is actually turning one into a narcissistic, self-centered maniac who is only obsessed with

¹⁵¹ See Kristjánsson 2020 on a neo-Aristotelian account of flourishing in education.

¹⁵² An important Aristotelian element is that one needs the external conditions in order to be in the position to flourish.

¹⁵³ As some forms of self-help may call for, see Hämäläinen 2013.

personal goals and moral self-training. One may believe that flourishing amounts to being happy,¹⁵⁴ yet discover that happiness is overrated. One may believe that flourishing amounts to devoting one's resources to altruistic causes,¹⁵⁵ yet discover that such causes root deeply into structural problems that should instead be the main focus of one's moral action. One may think that what flourishing amounts to may be found in scientific research about the human psyche,¹⁵⁶ yet discover that one need just to pay attention to humans in order to understand what it would be for them to flourish. An open-ended account of flourishing takes seriously the naturalistic claim that there is something that it is for humans to flourish, but does not tie it to universalistic, inflexible accounts of any sort. The normative prescription of an open-ended flourishing is, thus, something like the following: you should be in the position to X and you should X, until you discover that X is not all there is to it – because you *will* discover it. One of course need to be epistemically well-equipped in order to be able to find action-guidance and, at the same time, be open to the possibility that that action-guidance may be transcended, or undergo a paradigm shift. In what follows I propose a by no means exhaustive list of epistemic virtues that, if the situation requires, may be transformative to the extent to which they are likely to enable moral paradigm shifts and successful readjustments. I begin by expanding on epistemic virtues, then I consider those that I find most apt to do the job of construing a progressive moral outlook.

¹⁵⁴ As a result of the success that positive psychology had had on our everyday choices, see Cabanas and Illouz 2019.

¹⁵⁵ As an effective altruist may argue, see Syme 2019 for an overview and an objection.

¹⁵⁶ See VanderWeele's measure of flourishing (2017), applied by The Human Flourishing Program at Harvard's Institute for Quantitative Social Sciences, <https://hfh.fas.harvard.edu/measuring-flourishing>.

3.4 Two Examples

In what follows I consider two examples of MPS in relation to individual choices that are linked with current controversial, or difficult to navigate, issues. The aim is not that of arguing that everyone should act the way the people in the examples do, nor to show that these issues are actually easy to address. Rather, the aim is that of portraying possible ways in which a progressive moral outlook can frame these issues, highlighting the role of transformative virtues and of the new blueprint of flourishing that these fictional people come to develop. Ultimately, what I wish to stress is the fact that a VE progressive moral outlook need not be continuous with traditional moral paradigms, nor true to traditional biologicistic ways to understand human flourishing.

Parenthood

Kia has always valued being a mother and has built her life around making sure to be in the economic position to have a big family. One day Kia is walking down the street and sees a mother having a hard time with her crying child. She looks burnout; she tries to calm the child, while at the same time avoid creating a jam on the street, while also trying to figure out what to do with what she had already planned. Exhausted, she just waits for the child to stop crying, standing hopelessly in a corner, passed by annoyed pedestrians who are anxious to get somewhere.

Kia has always felt empathy for children: she loves them, enjoys listen to their high-pitch funny voices, and finds them absolutely adorable even when they are being naughty. Every time she has seen a kid crying for some desires to be immediately met, she has always thought that her parents must have done something wrong, and that she would have never found herself in the same position, since she would have known how

to educate her kids. However, in that moment she experiences something new: a strong, unapologetic empathy for the mother. Just by looking at her, she cannot help but see that no, she has not done anything wrong, she is just an exhausted human being who sometimes also needs time for herself, so that she does not give up her entire beings in order to be just that – a mother. The mother's reality is now disclosed and Kia can clearly see loneliness and exhaustion, rather than personal fulfillment. She feels uneasy. Her vision of motherhood is suddenly not consistent with valuing motherhood above other things anymore. She looks at her now obsolete moral paradigm, according to which she ought to have kids because it is just the way it has to be, and does not find it action-guiding anymore. Now she needs to construct a new one, starting from what it is that she has *seen* in virtue of the Gestalt switch crying child/lonely mother. Does having children involve giving them the whole of you? Is it right to be so? Do I want to? Are there other possibilities in life I am not considering? Who said that this is what makes women's life flourishing?

In this murkiness, *phronesis* allows her to take seriously her new intuitions about flourishing and integrate them with her new affordances: it is true that kids need full attention from adults, and that they need adults to be there for them, to take them seriously, to be supportive and solid – in short, that adults *ought to* be there for children. On the other hand it is also true that kids need *adults*, not necessarily their biological *mothers*, to be there for them. Kia has never considered this subtle difference before. Virtues of intellectual wholeness allow her to scrutinize where she comes from and see that her family's moral paradigm has always considered family in the strictly traditional sense; that is, biological family, where mothers are the only ones who naturally know what is best, and who draw meaning for their lives by taking care of children. She is in the position to question all that now. Her mental creativity and her intellectual transparency allow her to engage in an honest conversation with her partner and come to

an innovative conclusion. Her new moral paradigm has taken shape: parenthood can now be a shared quest, not her own private burden as a woman. She realizes that her best friend and her partner would like to have kids too, so she proposes to form a family nucleus with them, where they could be four legal parents of any number of kids they wish to give birth to, or adopt.

Charity

Oscar regularly gives to charities that provide food and material goods to the poor. He does that because he values effective altruism above everything else. Since he was a kid, his educators have taught him that he should make a fortune, so that he could be in the position to give to the worse well-off, and actually do something concrete for them, something effective. He is now rich enough to live a more than comfortable life, grow savings and give to charity. So he does that – he ought to.

One day he is sitting at a café working on his laptop, when all of a sudden he finds himself witnessing a chase: the waiter is running after a man who has stolen a sandwich. The man who is being chased comes towards the direction of Oscar; he is in his forties, wears dirty, worn-out clothes and is clearly malnourished. Oscar feels repulsion, but tells himself “don’t be mean, he’s just poor.” He stands up and waves some cash at the waiter: “leave him be, I’ll pay for his sandwich” says Oscar. The waiter takes the cash and goes back to the counter. Oscar turns to the poor man and prepares to say “it’s ok, really, don’t mention it. Is there anything else I can get you?” The poor man walks towards Oscar – who holds his breath while maintaining his humble expression – and slams the sandwich on the table. He gives Oscar a long, hateful stare, then exits the café grounds without his lunch. “So ungrateful!” people comment, sharing Oscar’s

puzzlement and irritation. “He probably wasn’t that hungry after all”; “he just wanted to steal,” comments go on. Oscar, however, after a moment of resentment, begins to feel something different. In the man’s hateful gaze, he saw something else, something utterly new and devastating: there was much more dignity in stealing a sandwich than in being helped by a wealthy philanthropist. The old paradigm according to which effective altruism was unquestionably among Oscar’s core values is now shattered. The new paradigm is constructed on the new affordances he has seen: the poor man suffers from a kind of injustice that is way beyond the mere not being able to afford basic commodities. The poor man has been stripped away of his dignity qua human being, and it is not by effectively giving him material goods that Oscar can give the man his dignity back.

This new normative reality is revealed through phronesis’ perceptual function; virtues of intellectual wholeness allow Oscar to question his previous paradigm, and to truly experience the cognitive dissonance needed in order to disrupt it. Virtues of mental flexibility now enter the picture: Oscar needs another, radically new moral paradigm in order to know what to do from now on. Finally, epistemic justice allows Oscar to consider that he does not know anything about being poor, about what it is to be stripped away of dignity. His new moral paradigm would consider this ignorance on his part, and would give proper recognition to the man’s reaction. We can imagine Oscar going back to the man’s expression over and over again: his eyes were saying “it’s your fault, don’t you dare buy food for me, don’t you dare show off how good you are, don’t you dare pretend we are the same, that you are just helping a fellow human being. I’m here because of you.” We can also imagine Oscar being puzzled by the accusation in the man’s eyes, but still thinking that there must be something he is missing, because he does not belong to the man’s world. Soon, he comes to realize that the problem is much more structural than he thought: he reads journal articles about the charity he donates to

and comes to learn that, in fact, this charity is making money out of the very existence of poor people; therefore, the interest is not much that of eradicating poverty in the neighborhood, but rather that of keeping it under control. Nobody should die of starvation, yet nobody should truly exist poverty. Oscar's moral outlook has shifted from being self-centered – i.e. “I want to feel good about the world being full of poverty” – to other-centered – i.e. “I truly wish nobody was stripped out of dignity” –; a shift that may not have happened if the man had not stared at him expressing all that with his eyes. He begins to look for solutions at the individual level – what it is that *he* could do – that also, and above all, consider the structural level – how to reach such a poverty-free society.

Once again, depicting these two examples was not aimed at arguing what it is that one should ultimately do; that is, I do not know better than Kia and Oscar. The point was rather to picture what a progressive moral outlook may look like, and how it can help addressing pressing current issues that may seem scary precisely because they are current: that is, they require new moral paradigms to be made sense of.¹⁵⁷

The concept of “parenthood” is quickly changing, and new generations of adults are likely to be unequipped to deal with these changes. Is being a parent a *good* choice, given the current circumstances? Is it environmentally friendly? Does becoming a mother automatically relegate women to unpaid care work? If so, are we still ok with this? Is it selfish to want children? Is it selfish *not* to want to have children? And again, parenthood is strictly related to the concept of “family,” which is also quickly changing. The centrality of biological relations in families are finally questioned by transfeminist

¹⁵⁷ What I have depicted here can be applied to new, potentially scary issues such as the growing technology behind AI and pressing changes of direction when it comes to the environment, which call for instant action.

and LGBTQI+ advocates, who stress the fact that parenthood is about love and not about blood relations.¹⁵⁸ How should future families look like? A conservative moral outlook would search for answers in existing moral paradigms; for this reason, they will never be able to equip new generations to be able to make their own, conscious choices on these matters. A progressive moral outlook considers the fact that the future of families is open-ended – that is, there will be other moral paradigms that may be more accurate than the one we may construct now, because situations change quickly – while at the same time allow current generations of adults to have new, innovative conceptual tools to handle the hard questions; and all this can be granted, while also considering the structural problems that support the perpetration of injustices related to those issues, such as women’s relegation to unpaid care work, the difficulty – or the impossibility – for queer parents to have children, the effort to maintain women’s choices on motherhood as dichotomies between pursuing a career or having children, and many more.

Oscar’s example is an intricate net where we can distinguish structural problems proper of capitalist societies, where social classes are on the one hand rigidly maintained, while on the other are said not to exist, because if only one wanted to, one could exist poverty anytime, and where charity institutions are thus the pillar of this double standard on poverty.¹⁵⁹ From that, we can also distinguish the individual choices of a rich man who is the product of this society and who endorses libertarian values such as freedom to help the way one wants to, e.g. by giving money to charity rather than paying taxes.

¹⁵⁸ See any of latest issues on Taylor & Francis’ Journal “LGBTQ+ Family”.

¹⁵⁹ This example owes its narration to Oscar Wilde’s 1891 “The Soul of Man under Socialism”, where he provides a critique of libertarian values of capitalist societies and considers the very existence of charities and the stress on developing altruistic values, as *immoral* precisely because they purely aim at maintaining the capitalist, deeply unequal status quo.

If one truly valued the eradication of poverty, what should one do? How could one actually *do* something about such a structural level problem, if the solution is not proven to be as simple as “give money”? And again, how does the existence of world poverty impact on our everyday consumerist choices? Or on our career choices? These questions are not new in and of themselves. However, current moral paradigms have proven to be inadequate to even conceive of this problem in both an accurate and action-guiding manner. Thus, it is pressing to find a new, progressive one; that is, one that disrupts current value systems that frame this problem, e.g. that of looking at poverty through the lenses of meritocracy, thus thinking that one is poor because one deserves to be so, or through the lenses of effective altruism, thus thinking that poor people are just unlucky and we lucky ones should effectively help them through donations. That is, giving them the crumbs that fall off our dinner tables.

4. Conclusion

In this chapter I started from the following observation: MPS are progressive changes to the extent to which they enable agents to transcend existing moral paradigms and adopt new, more accurate ones. Thus, a moral theory that accounts for the possibility for agents to undergo MPS is a progressive moral theory, since it accounts for the possibility for agents to adopt new, more accurate moral paradigms. I then considered MPS within the framework of neo-Aristotelian Virtue Ethics and stressed the fact that such framework does not in and of itself account for the possibility that agents undergo MPS. Thus, it is inherently conservative.

I then isolated the problem of conservatism that is proper of Virtue Ethical methodology and explored how virtues are generally conceived: one that sees virtues as originating in moral traditions (MacIntyre), the other that sees virtues as originating in facts about human nature (Foot) and noticed that while the former is hopelessly conservative, the latter need not be. Or rather, in the form advocated by Foot it is; but in the form advocated by McDowell – that is, second nature naturalism – it need not be. I then defended an account of second nature naturalism that finds its origins in Murdoch's idea of moral progress as attuning to normative reality but, with Severini, without the necessity to stipulate the existence of a metaphysical reality to be in the position to make progressive claims.

In the third and last section I argued that if neo-Aristotelian VE can integrate moral development with epistemic development, then it could be equipped to make room for the possibility of MPS and, thus, be a progressive rather than a conservative moral theory. In particular, I argued that it can and that it should do so. I then proposed a by no means exhaustive list of epistemic virtues that, if developed, could prove to be “transformative” to the extent to which would go to construe a moral outlook that is

capable of dealing with MPS; namely, phronesis, virtues of intellectual wholeness (such as intellectual integrity, honesty, humility, transparency, self-awareness and self-scrutiny), virtues of intellectual flexibility (such as imaginativeness, creativity, intellectual flexibility, open-mindedness, agility and adaptability), and virtues of epistemic justice in both their testimonial and hermeneutical declinations.

I concluded with two practical examples of how a progressive moral outlook would look like and would work through radical moral transformations when it comes to current, pressing moral issues that call for our attention and, in particular, for new, innovative ways to be framed.

References

Annas, J. (2011). *Intelligent Virtue* (Oxford, Oxford University Press).

Andreou, C. (2006). 'Getting on in a Varied World', *Social Theory and Practice* 32.1: 61-73, 10.5840/soctheorpract20063213

Anscombe, G.E.M. (1958). "Modern Moral Philosophy," *Philosophy*, 33: 1–19.

Appiah, K.A. (2010). *The Honor Code: how Moral Revolutions Happen*. W. W. Norton & Company.

Aristotele (1999) *Etica Nicomachea*, It. tr. Natali Carlo (a cura di), Edizioni Laterza, Bari.

Aristotle. (1905). *Politics*. Oxford: Clarendon Press,

Baehr, J. (2011). *The Inquiring Mind: On Intellectual Virtues and Virtue Epistemology*, Oxford University Press.

Battaly, H. (2015). *Virtue. Polity*.

Beiser, F., (1992). *Enlightenment, Revolution and Romanticism: The Genesis of Modern German Political Thought 1790–1800*, Harvard: Harvard University Press.

Brewer, T. (2009). *The retrieval of ethics*. New York: Oxford University Press.

Cabanas, E., Illouz, E. (2019). *Happycracy. Come la scienza della felicità controlla le nostre vite* (original title: *Happycracy. How the science of happiness controls our lives*). Codice Edizioni.

Caprioglio Panizza, S. (2022). *The Ethics of Attention. Engaging the Real with Iris Murdoch and Simone Weil*. New York: Routledge.

Carr, D., Steutel, J. (eds.), (1999). *Virtue Ethics and Moral Education*, New York: Routledge.

Chappell, S. G. (2014). *Knowing What to Do*, Oxford: Oxford University Press.

Croce, M. (2020). *Teorie dei Vizi. Un'analisi critica << Etica & Politica >> 22: 577-598.*

Croce, M., Liberti, M., Vaccarezza, M. S. (2024). *The Power of Fame: Celebrities and Influencers as Moral Exemplars*, in Robb, Catherine M.; Archer, Alfred & Dennis, Matthew (eds.) (2024). *Philosophy of Fame and Celebrity*. Bloomsbury.

Curzer, H. J. (1998). *To Become Good*. Paper presented at the 20th World Congress of Philosophy, Boston (MA), 10–15 August 1998. Available at: www.bu.edu/wcp/index.html, accessed 26 April 2012.

Curzer, H. J. (2002). *Aristotle's Painful Path to Virtue*, *Journal of History of Philosophy*, 40.2, pp. 141–162.

Curzer, H. J. (2012). *Aristotle and the Virtues* (Oxford, Oxford University Press).

Dancy, J. (2006). Nonnaturalism. In David Copp (ed.), *The Oxford handbook of ethical theory*. New York: Oxford University Press.

Darnell, C., Fowers, B. J., and Kristjánsson, K. (2022). A multifunction approach to assessing Aristotelian phronesis (practical wisdom). *Personality and Individual Differences* 196:111684.

De Caro, M. Vaccarezza, M. S. & Niccoli, A. (2018). Phronesis as Ethical Expertise: Naturalism of Second Nature and the Unity of Virtue. *Journal of Value Inquiry* 52 (3):287-305.

De Caro, M., Bina, F., Bonicalzi, S., Brunetti, R., Croce, M., Kerasausaite, S., Navarini, C., Ricci, E., Vaccarezza, M. S., forthcoming in *The Journal of Medicine and Philosophy*.

Driver, J. (2003). The Conflation of Moral and Epistemic Virtue. *Metaphilosophy* 34 (3):367-383.

Eriksson K., Coultas J.C. (2014). Corpses, maggots, poodles and rats: emotional selection operating in three phases of Cultural Transmission of Urban Legends. *J Cognition Cult* 14(1–2):1–26. <https://doi.org/10.1163/15685373-12342107>.

Fine, K. (2002). Varieties of Necessity. In Tamar Gendler & John Hawthorne (eds.), *Conceivability and Possibility*. New York: Oxford University Press. pp. 253-281.

FitzPatrick, W. J. (2008). Robust ethical realism, non-naturalism, and normativity. *Oxford Studies in Metaethics* 3:159-205.

Foot, P. (2001). *Natural Goodness*, Oxford: Clarendon Press.

Fricker, M. (2007). *Epistemic injustice: power and the ethics of knowing*. New York: Oxford University Press.

Garrett, J. E. (1993) The Moral Status of 'the Many' in Aristotle, *Journal of the History of Philosophy*, 31.2, pp. 171–189.

Goldman, A. (1998). "Reliabilism" in Edward Craig (ed.) *Routledge Encyclopedia of Philosophy*.

Golob, S. (2014). *Heidegger on Concepts, Freedom and Normativity*. Cambridge: Cambridge University Press.

Greco, J. (2000). Two Kinds of Intellectual Virtue <<Philosophy and Phenomenological Research>> 60: 179-184.

Hacker-Wright, J. (2021). "Philippa Foot", *The Stanford Encyclopedia of Philosophy* (Winter 2021 Edition), Edward N. Zalta (ed.), URL = <<https://plato.stanford.edu/archives/win2021/entries/philippa-foot/>>.

Hämäläinen, N. (2016). Self-Help, Moral Philosophy, and the Moral Present. *Human Studies* 39 (2):289-306.

Hamilton, A. (2020). "Conservatism", *The Stanford Encyclopedia of Philosophy* (Spring 2020 Edition), Edward N. Zalta (ed.), URL = [<https://plato.stanford.edu/archives/spr2020/entries/conservatism/>](https://plato.stanford.edu/archives/spr2020/entries/conservatism/).

Heidegger, M. (1927). *Being and Time*, trans. John Macquarrie and Edward Robinson (New York: Harper, 1962), foreword by Macquarrie and Robinson (xiv).

Hermann, J. (2017). Possibilities of Moral Progress in the Face of Evolution. *Ethical Theory and Moral Practice* 20(1):39–54. <https://doi.org/10.1007/s10677-016-9737-2>.

Hobsbawm, E. and T. Ranger (eds.), (1983). *The Invention of Tradition*, Cambridge: Cambridge University Press.

Hookway, C. (2003). How to Be a Virtue Epistemologist in DePaul and Zagzebski (eds.), pp. 146-160.

Hu, J.I. (2021). Shame, vulnerability, and change. *J Am Philosophical Association* 1–18. <https://doi.org/10.1017/apa.2021.21>.

Jubilee Centre for Character and Virtues (2022), *The Jubilee Centre Framework for Character Education in Schools*, Birmingham: University of Birmingham [Online], Available at: <https://www.jubileecentre.ac.uk/userfiles/jubileecentre/pdf/character-education/Framework%20for%20Character%20Education.pdf>

Kidd, J. (2021). A Case for an Historical Vice Epistemology, <<HumanaMente>> 14 (39), 69-86.

Kidd, Ian James; Cassam, Quassim & Battaly, Heather (eds.) (2020). Vice Epistemology. New York, NY: Routledge.

Knight, K. (1996). Revolutionary Aristotelianism, in Virtue and Politics. Alasdair MacIntyre's Revolutionary Aristotelianism 2011.

Kotsonis, A. (2021). The Aristotelian understanding of intellectual vice: Its significance for contemporary vice epistemology. European Journal of Philosophy 30 (3):1161-1172.

Kristjánsson, K. (1998). Liberating moral traditions: Saga morality and Aristotle's megalopsychia. Ethical Theory and Moral Practice 1 (4):397-422.

Kristjánsson, K. (2007). Aristotle, Emotions, and Education (Farnham, Ashgate).

Kristjánsson, K. (2010). The Self and its Emotions (Cambridge, Cambridge University Press).

Kristjánsson, K. (2015). Aristotelian Character Education. Routledge.

Kristjánsson, K. (2020). Flourishing as the aim of education: a neo-Aristotelian view. New York, NY: Routledge/Taylor & Francis Group.

Kuna, M. (2005). Macintyre on tradition, rationality, and relativism. *Res Publica* 11 (3):251-273.

LeBar, M. (2013). *The Value of Living Well*, New York: Oxford University Press.

Liberti, M. (2022). Virtue and Continence: Defending their Cognitive Difference. *Philosophical Inquiries* 10 (2):39-58.

Lott, M. (2020). ‘Aristotelian Naturalism and the Autonomy of Ethics’ in M. Hähnel (ed.), *Aristotelian Naturalism: A Research Companion*, Springer Nature.

MacIntyre, A. (1985). *After Virtue*, London: Duckworth, 2nd Edition.

Maguire, B. (2015). *Grounding the Autonomy of Ethics*. *Oxford Studies in Metaethics* 10.

McDowell, J. (1998). “Two Forms of Naturalism”, in *Mind, Value, and Reality*, Cambridge, MA: Harvard University Press.

Millgram, E. (2009). ‘Life and Action’, *Analysis* 69: 557–564, 10.1093/analys/anp087

Moosavi, P. (2022). “Neo-Aristotelian naturalism as ethical naturalism”, *Journal of Moral Philosophy*, 1(aop): 1–26.

Morrow, P. (2020). Identity-Directed Norm Transformations and Moral Progress. *J Value Inq* 54(3):493–509. <https://doi.org/10.1007/s10790-019-09722-4>.

Murdoch, I. (1971). *The Sovereignty of Good*, London: Routledge.

Nichols, S. (2002). On the genealogy of norms: a Case for the role of emotion in Cultural Evolution. *Philos Sci* 69(2):234–255. <https://doi.org/10.1086/341051>.

Ohlhorst, J. (2023): Engineering virtue: constructionist virtue ethics, *Inquiry*, DOI: 10.1080/0020174X.2023.2238293.

Plato. (1943). *The Republic*. New York: Books, Inc.

Rehren, P., & Sauer, H. (2022). Another Brick in the Wall? Moral Education, Social Learning, and Moral Progress. *Ethical Theory and Moral Practice*, 1-16.

Rosen, G. (2017). “Metaphysical Relations in Metaethics”, in Tristram McPherson and David Plunkett (eds.), *The Routledge Handbook of Metaethics*, New York: Routledge, 151–169.

Sanderse, W. (2015). An Aristotelian Model of Moral Development. *Journal of Philosophy of Education* 49 (3): 382-398.

Sauer, H. (2019). Butchering Benevolence Moral Progress beyond the Expanding Circle. *Ethical Theory and Moral Practice* 22(1):153–167. <https://doi.org/10.1007/s10677-019-09983-9>.

Severini, E. (2021). Moral Progress and Evolution: Knowledge Versus Understanding. *Ethical Theory and Moral Practice*, 24(1), 87-105.

Sherman, N. (1989). *The Fabric of Character. Aristotle's Theory of Virtue* (Oxford, Clarendon Press).

Silverstein, A. and Trombetti, I. (2013). Aristotle's Account of Moral Development, *Journal of Theoretical and Philosophical Psychology*, 33.4, pp. 233–252.

Slote, M. (1995). Agent-Based Virtue Ethics. *Midwest Studies in Philosophy* 20 (1):83-101.

Slote, M. (1997). "Virtue Ethics", in Marcia Baron, Philip Pettit, and Michael Slote, *Three Methods of Ethics*, Oxford: Blackwell, pp. 175–238.

Slote, M. (2001). *Morals from Motives*, Oxford: Oxford University Press.

Snow, N. E. (2018). 'Neo-Aristotelian Virtue Ethics', in Nancy E. Snow (ed.), *The Oxford Handbook of Virtue*, Oxford Handbooks (online edn, Oxford Academic, 6 Dec. 2017), <https://doi.org/10.1093/oxfordhb/9780199385195.013.34>

Sosa, E. (1980). The Raft and the Pyramid: Coherence versus Foundations in the Theory of Knowledge, <<Midwest Studies in Philosophy>> 5: 3-25.

Sturgeon, N. (2002). Ethical Intuitionism and Ethical Naturalism. In Philip Stratton-Lake (ed.), *Ethical Intuitionism: Re-Evaluations*. Oxford University Press UK.

Swanton, C. (2001). A virtue ethical account of right action. *Ethics* 112 (1):32-52.

Swanton, C. (2003). *Virtue Ethics: A Pluralistic View*, Oxford: Oxford University Press.

Swanton, C. (2021). *Target Centred Virtue Ethics*, Oxford: Oxford University Press.

Swartwood, J. (2013). Wisdom as an expert skill. *Ethical Theory and Moral Practice* 16(3):511-28.

Syme, T. (2019). Charity vs. Revolution: Effective Altruism and the Systemic Change Objection. *Ethical Theory and Moral Practice* 22 (1):93-120.

Tanesini, A. (2021a). *The Mismeasure of the Self. A study in Vice Epistemology*. Oxford University Press.

Tanesini, A. (2021b). Ignorance, arrogance, and privilege: vice epistemology and the epistemology of ignorance, in Kidd, I. J., Battaly, H. and Cassam, Q. (eds) *Vice Epistemology*, Routledge.

Taylor, C. (2006). "Engaged Agency and Background in Heidegger," in Charles B. Guignon (ed.), *The Cambridge Companion to Heidegger*, 202–21.

Tobin, B. (1989) An Aristotelian Theory of Moral Development, *Journal of Philosophy of Education*, 23.2, pp. 195–211.

Tsai, C. (2022). *Wisdom: A Skill Theory*. Cambridge: Cambridge University Press.

VanderWeele, T.J. (2017). On the promotion of human flourishing. *Proceedings of the National Academy of Sciences, U.S.A.*, 31:8148-8156.

Viaggiano, A., Galletti, M. (2007). ‘George Edward Moore e il dibattito sul naturalismo’, in G. Bongiovanni (a cura di) (2007). *Oggettività e morale. La riflessione etica del Novecento*. Milano: Bruno Mondadori.

Wilde, O. (1891) *L’anima dell’uomo sotto il socialismo*. Universale Economica Feltrinelli, sesta edizione (2022).

Wilson, C. (2010). Moral Progress without moral realism. *Philos Pap* 39(1):97–116.

Wolt, D. (2022). Phronêsis and Kalokagathia in Eudemian Ethics VIII.3. *Journal of the History of Philosophy* 60 (1):1-23.

Woodcock, S. (2006). ‘Philippa Foot’s Virtue Ethics Has an Achilles’ Heel’, *Dialogue* 45: 445–468, 10.1017/S0012217300001013

Wrathall, M. A. (2011). “Unconcealment” in Mark A. Wrathall, *Heidegger and Unconcealment: Truth Language and History*. New York: Cambridge University Press, 11–39, 15.

Zagzebski, L. (1996). *Virtues of the Mind*, New York: Cambridge University Press.

Zagzebski, L. (1998). "The Virtues of God and the Foundations of Ethics," *Faith and Philosophy*, 15 (4): 538–553.

Zagzebski, L. (2004). *Divine Motivation Theory*, New York: Cambridge University Press.

Conclusion

People do change, and the way they can do so takes different forms. In Chapter I we explored three main conceptualizations of radical moral change – moral conversion, moral transformation and moral progress – and drew some preliminary observations on both commonalities and points of departure. In Chapter II I proposed three parameters to help distinguish among different forms of moral change – (i) whether the change is active or passive, (ii) whether the change is experienced as gradual or abrupt, and (iii) whether the change is core or peripheral – and isolated one particular form that I believed worth of attention, which is that of changes that are passive, abrupt and core (PAC). I then drew a parallel between PAC changes and Kuhn’s paradigm shifts in science and argued that PAC changes were of the same form. Thus, I proposed to call them “Moral Paradigm Shifts” (MPS). In Chapter III I expanded on the importance of making sense of MPS: since they are changes in which an existing paradigm is replaced by a new, more accurate one – as it is the case for scientific paradigm shifts – then MPS are progressive changes; in other words, they are instances of *real*, Wittgensteinian moral progress. I then advocated for a type of neo-Aristotelian Virtue Ethics that could account for the possibility for agents to undergo MPS and, thus, avoid the charge of conservatism. Progressive Virtue Ethics necessarily integrates moral development with epistemic development: that is the key, I argued, to ensure that account of “flourishing” can be both normative and dynamic; that is, at least in principle a possible subject of MPS.

The underlying thread is the following: we are urged by current times to change the way we frame morally significant issues, such as how we relate with nature, what types of relationships we build with other humans and, why not, with the artificial intelligence that talks to us from our devices. The hope is that we are able to look at the same moral

phenomenon we may find scary – e.g., the end of the “traditional” family as a relationship standard – or impossible to address – e.g., climate change – and see something we have not seen before; a moral reality that can guide us towards progress, in virtue of its being a more accurate way of framing the problem. Our moral theories should equip us to be at least in principle able to undergo MPS, and Virtue Ethics can in fact do so.