Di Francesco, Michele, Marraffa, Massimo and Paternoster, Alfredo, *The Self and Its Defences. From Psychodynamics to Cognitive Science.*
London: Palgrave MacMillan, 2016, pp. vii + 219.

Today more than ever before, the philosophical debate on the self is situated at the crossroads of diverse disciplines.[1] The multifaceted and integrative nature of the contemporary researches on subjectivity is the first noticeable aspect widely rendered in *The Self and Its Defences. From Psychodynamics to Cognitive Science,* where the Authors M. Di Francesco, M. Marraffa and A. Paternoster outline their theory of the self and self-consciousness. Framed in a naturalistic as well as theoretically informed scenario, the theory they propose stems from the puzzling attempt to define what is a self, in the light of a conceptual and empirical inquiry, along with an historical and analytical reconnaissance of some relevant issues in philosophy of psychology and philosophy of mind.

From a philosophical point of view, the reflection on the self and self-consciousness is currently located between apparently opposed and yet complementary and sound conceptual options, whose differing accounts mirrors to a certain extent the common sense twofold intuition. Indeed, on the one hand, the self appears undoubtedly a product of the cultural realm, something constructed through social interaction, communication and narratives, i.e. the elemental, psychological unit of the broader, juridical concept of *person*. On the other hand, a self seems to be the starting condition we need in order to consider an organism a subject of experience, a psychobiological system bounded in his first-person perspective, somehow emerging from neural and biological mechanisms. It can easily be expected though, that even common sense is reluctant to define the self as merely an abstract entity constructed through inferences and interpretations; or conversely to think that it merges entirely with the primordial, bodily and perceptual domain of the first-person perspective; or farther that it overlaps with the mechanisms underlying the personal domain. *The Self and Its Defences* aims at combining these constitutive aspects emerging from the philosophical and common sense reflection, and at maintaining both a *top-down* and a *bottom-up* explanation, in the wake of the challenge promoted by cognitive science, which consists in "widening the naturalistic realm […], denying that our choices and actions can be exhaustively attributed to historical and interpretative factors, and thus trying to overcome the dichotomy, dear to the hermeneutic tradition, between *Naturwissenschaft* and *Geisteswissenschaft*" (11).

In the domain of the contemporary debate on subjectivity, the project of making this connection possible is relevant and worthy, as in the last decade a certain amount of philosophical works has dealt with the abovementioned alternative.[2]

Non reductive, explanatory pluralism is the compass of the project endorsed in the book, where philosophy of psychology constantly dialogues with

[1] See Gallagher, S. (ed.) 2011, *The Oxford Handbook of The Self,* Oxford: Oxford University Press.
[2] See Gallagher, S. 2000, "Philosophical Conceptions of The Self: Implications for Cognitive Science", *Trends in Cognitive Science*, 4, 1, 14-21; Zahavi, D. 2007, *Self and Other: The Limits of a Narrative Understanding*, in D.D. Hutto (ed.), *Narrative and Understanding Person*, Royal Institute of Philosophy Supplement, 60, Cambridge: Cambridge University Press.

several domains that ranges from philosophy and metaphysics of mind, to phenomenology, anthropology, developmental and social psychology.

The derivative approach to the self is usually connected to the philosophical thesis of anti-realism, that is, whereas the self is not considered as something primitive and logically prior, or as just a bodily, bounded entity that perceives the world from its first-person perspective, but rather as a product, it ends up being an illusion or a useful fiction, a work of art, namely an entity which is not part of the equipment of the natural, physical world. But a consideration of the self in terms of solely neurocognitive, bodily processes underlying the personal domain determines a reductive approach as well. More specifically, the effort to outline a naturalistic scenario where elusive notions like that of self can be included has to face the well-known trouble of the overlapping between the personal and the automatic, blind, sub-personal domain. Historically, the attempt to combine naturalism and realism when treating the abstract entities posited by psychology lies at the very ground of almost every debate in philosophy of mind. As of the more specific debate on the self, this effort arises from the overcoming of an old, idealistic approach, back to Descartes' *Cogito* to the Kantian *I Think* and the Husserlian transcendental phenomenology. *The Self and Its Defences* continues and renews the philosophical tradition that has questioned the given, innate nature of the self, and combines this tradition with the conceptual framework of the contemporary, empirical researches on subjectivity. The sound intuitions and the large empirical data that support a naturalistic account of the self are explored and evaluated without accomplishing a reduction to the sub-personal domain.

The Authors state at the very beginning that the theory of the self they outline is at the same time a theory of self-consciousness, declaring a pivotal idea developed in the book. Indeed, the connection between the two notions (i.e., self and self-consciousness) reveals not only that the self is considered in its reflexive and "objectifying" component, i.e. in its being both a subject and an object of consciousness. It also establishes that in order for an organism to be considered a self, it has to be involved in a process of knowledge, namely of transitive consciousness, where the object of consciousness is the self. The two core ideas are that the self is "a psychobiological system activity of self-representing", and that this representational process is aimed at "defending the self-conscious subject against the threat of its metaphysical inconsistence" (1). Following William James, the theory here proposed is focused on the mechanisms that allow the I to make the Me, i.e. on both the processual, representational activity of the mind, and the representation of the self as an object of consciousness. Chapter 2 (*The Unconscious Mind*) and Ch. 3 (*Making the Self, I: Bodily Self-Consciousness*) are mainly dedicated to the elaboration of the first thesis, i.e. to the description of the self's processual, representational activity, starting from the unconscious mechanisms to the gradual emergence of object consciousness. In Chapter 4 (*Making the Self, II: Psychological Self-Consciousness)* and Ch. 5 (*The Self as a Causal Centre of Gravity*), it is presented the second core idea of the theory, i.e. the defensive nature of the self-*ing* process and the more general theme of the fragility of the subject.

The structure of the book mirrors the bottom-up route aimed at individuating the several stages towards the construction of personal identity. After having introduced the methodological and conceptual frame in which their research is conducted (Ch. 1), the Authors present an analytical and historical examination

of the notion of the unconscious, which lies at the very foundation of the subjectivity's development (Ch. 2). The core of this section is that the abilities manifested in behaviour have to be explained by the unconscious mechanisms underlying them. This thesis calls into question the puzzling mark of the mental problem (i.e., how are we to define the boundaries between what is mental and what is not?), the connected interface problem (i.e., how are we to combine and make the personal level of explanation interacting with the non-personal one?), as well as an inquiry on the difference between the Freudian and the psychodynamic unconscious. As of the first issue, while trying to individuate the right criterion for demarcating the mental, it is auspicated a strategy for the indirect incorporation of the sub-personal as well as its dialectical relationship with the personal level. As of the second issue, in overcoming both Freud's positivistic naturalism and the hermeneutical, post-modern reading of psychoanalysis, cognitive science is situated halfway between the personal sphere of the first-person phenomenology and the non-personal domain of neurobiological events. The historical inquiry of this section brings out a fundamental issue: if the cognitive unconscious, far from reflecting the conscious level, is not mental, how to interface the computational level of explanation with the ordinary psychological explanations? What it is auspicated is a coexistence and compatibility between the ordinary image of the subject as a rational, free agent, and the scientific conception of subjects as computational machines, in a strategy where both dependence and autonomy are worth to be maintained. In this perspective, the concept of attachment analysed in a cognitive and evolutionary framework is taken to be a pivotal research area in which the relationship is possible. The rehabilitation of psychoanalysis in a cognitive-evolutionary framework, where dynamic psychology and the cognitive sciences are connected through the notions of motivation and attachment, determines a systemic–relational framework where neither radical social constructivism nor strong individualism are embraced. In this developmental and evolutionary framework, minds are shaped in early interactions with others, and personal identity's assembling is followed starting from the analysis of the young child's affective and cognitive relationality.

In Ch. 3 a non-idealist, non-eliminative view of the self is outlined. The section contains an analytical discussion of Humean eliminativism on the self, of analytic Kantianism, and of the contemporary phenomenological accounts of self-consciousness. This examination aims at combining naturalism and realism, thus at questioning the thesis according to which the self has no room in the natural world. The worth-mentioning insight of this section is that the reality of the self does not rest on a phenomenological, bodily and innate sense of *me-ness* or *for-me-ness*, which turns out to merge with a weak form of first-person perspective, or with a transcendental precondition evoking the Kantian *I Think*.[3] In the bottom-up construction of the self here endorsed, the early intelligent behaviour as well as the organism's experiential interaction with the environment and the boundedness of its body are not considered as developmental stages where a form of minimal, pre-reflective self-consciousness makes the experience possible. In particular, it is argued that in order to consider an organism self-conscious, it has to be conscious of its body taken as a whole, namely it has to be able to represent himself as the object of its consciousness. Therefore, it is

---

[3] See also Marraffa, M. and Paternoster, A. 2014, "A Third-person Approach to Self-Consciousness", *A&P, International Multidisciplinary Journal*, 11, 107-19.

questioned the very popular hypothesis that a form of pre-reflective self-consciousness grounds the reflective one as well as consciousness in general.[4] The pivotal thesis that emerges in this section, i.e. that self-consciousness does not ground conscious experience but is rather acquired during development, is then sustained by a large review of developmental psychology and cognitive ethology, and through a critical reading of some philosophical attempts to connect the a-priori methodological approaches on the study of subjectivity with the empirically informed ones. Besides the fact that the possession of a first-person perspective—with which the "empirically void" notion of pre-reflective self-consciousness seems finally to merge with—is not a sufficient condition for self-consciousness, the Authors finally argue that bodily self-consciousness is far from being consciousness of oneself as a continue entity through time, namely it is far from being a psychological self.

The emergence of an "extended"[5] self (Ch. 4) is due to an interplay of mentalization, memory and interpersonal skills regulated by cultural and environmental cues. The construction of the virtual inner space of the mind is followed within the framework of attachment theory, along with a systematic, remarkably large review of the literature ranging from developmental, to social and personality psychology. It is retraced the process towards the subject's awareness of the mind as an inner dimension, as well as its evolution into the cognitive more complex form of the narrative/autobiographical self.

The two core ideas underlying the theory outlined in *The Self and Its Defences*—i.e. the processual and defensive nature of the self-*ing* activity—are elaborated in a framework aimed at avoiding both idealism, according to which the self is somehow logically prior, and eliminativism, that conversely stems from a derivative account of the self. The naturalistic scenario that frames the book, far from embracing a consideration of the self as an illusion or an epiphenomenon, is then aimed at drawing a realistic account of the self. Through a strategy that contemplates the constant examination of bottom-up and top-down explanatory paths, the Authors' intent is to avoid the excesses of these approaches, that is, as of the latter, the reductive approach that stems from an analysis in terms of solely neurocognitive mechanisms, and as of the former, the inflationary and radical top-down approaches. Naturalism is endorsed through the systematic attention to the scientific, empirical researches of cognitive sciences, and to the contribution given by the attachment theory and the psychodynamic tradition.

Given the conciliatory project of this work, the demanding, fond-of-metaphysics reader could finally feel a bit unconvinced. At the very beginning, the Authors state that their project is more an exercise in the philosophy of psychology rather than in the metaphysics of mind. Nevertheless, the book is necessarily involved with fundamental issues such as the already mentioned mark of the mental problem, the causal closure of the physical domain, the overdetermination thesis, just to mention few controversial themes to which the Authors themselves allude. As of the mark of the mental problem, the metaphysical grounding is far from being insignificant. The sub-personal level is taken to be

---

[4] See Gallagher, S., Zahavi, D. 2005, "Phenomenological Approaches to Self-Consciousness", *The Stanford Encyclopedia of Philosophy* (Spring 2015 Edition), E.N. Zalta (ed.), http://plato.stanford.edu/archives/spr2015/entries/self-consciousness-phenomenological
[5] The term is drawn from Damasio, A. 1999, *The Feeling of What Happens. Body, Emotion and the Making of Consciousness,* London: Vintage.

connected to the personal one through the undeniable fact that the subject has access to the products of the former, but not to the sub-personal processes themselves (§ 1.3.2). What is mental, according to this proposal, is the integrated sub-personal realm, whose inclusion is due to the transparency of its products. This proposal rests on a metaphysical thesis, because the desired connection between the personal and the sub-personal level of explanation can be established only if sub-personal computations are considered as genuine pieces of mind: that is, just in case it is endorsed a metaphysical commitment to the extended mind hypothesis. Thus, the mark of the mental problem is an ontological, metaphysical problem, and unless the products of the sub-personal are considered as genuine pieces of mind, the purpose of connecting the sub-personal and the personal as much as it is possible hardly establishes what is mental and what is not from an ontological point of view.

As of the difficulty with the abovementioned conciliatory strategy, it lies in a crucial point, on which the Authors themselves take stock at the end of the book, and that could be a worth pursuing extension of their theory of the self. How does the narrativism they endorse differ from the hermeneutical one? And how does it avoid the illusionism *à la* Dennett?[6] As of the former, likely question, the framework in which their narrativism is maintained, they argue, is naturalistic because the theory-driven self-interpretation is taken to be a re-appropriation of the products of the neurocognitive unconscious. Second, it is argued that narrative selfhood would not arise without an affective and bodily self-description. Furthermore, the very reason why the Authors consider their theory a non eliminativist one lies in a sort of "final", double overcoming of the Cartesian account of the self. They observe how this account still echoes in contemporary eliminativism, for it seems to presuppose that if the self exists, it must be a persisting individual substance. In other words, according to the Authors, contemporary illusionism infers the non reality of the self from the lack of a neuronal counterpart of the Cartesian ego.

Another clarification aimed at pointing out the specificity of their theory is worth questioning. The Authors state that their disagreement with the eliminativists is not merely verbal, because they are not arguing that the illusion of the self can have causal powers as a false belief may (66). This is taken to be an "obvious" and "irrelevant" conceptual move (*Ibid.*), useless if considered in order to distinguish their account from the eliminativist one. Nevertheless, given the confabulatory, deceptive nature of the self here widely documented, the reader could ask why the definition of the self as a "causal center of gravity" (Ch. 5) should not be considered also in the sense in which a false belief has causal powers. After all, a crucial assumption is that the selfing process deals with a teleology of self-defense: far from being an epiphenomenon, subjective identity is a layer of personality, a causal center of gravity whose protection is necessary for mental health. As of this issue, a worth pursuing extension of the book could be a clarification of the criteria affording the distinction between self-knowledge and self-deception, in particular regarding their different causal role.

This distinction seems to depend on a fascinating criterion derived from developmental and evolutionary data: if, after all, the problem with antirealism is that it disregards the *defensive* nature of identity self-construction (147), does real-

---

[6] Dennett, D.C. 1992, "The Self as a Center of Narrative Gravity"*, in Kessel, F., Cole, P., Johnson, D. (eds.), *Self and Consciousness: Multiple Perspectives,* Hillsdale: Erlbaum.

ism finally rest on the adaptive benefit of the narrative self? If metaphysics is taken to be driven by our current best scientific theories, then the self is real as such theories seem to suggest that the very foundation of human beings' health is the protection from psychological disintegration, which is made possible by the production of causal efficacious' life stories.

University Roma Tre   MARIAFLAVIA CASCELLI

Williamson, Timothy, *Modal Logic as Metaphysics*.
Oxford: Oxford University Press, 2013, pp. xvi + 464.

At the end of the preface of *Modal Logic as Metaphysics* (MLAM, from now on), T. Williamson apologises (to his sons) for writing a book "with such a dull title" (p. xvi). Indeed, the title of this book does not even try to be engaging, or cool, but it has surely the merit of highlighting the fundamental methodological thesis on which the entire project is based. If modal logic *is* a metaphysics, then the role of modal logic within modal metaphysics—and that of logic within philosophy in general—is not (anymore) neither that of being a *neutral arbiter* nor—to use David Kaplan's words—that of "striving to serve ideologies".[1] This methodological attitude signs the end of *exceptionality* about logic and the beginning of an *anti-exceptional* attitude about it. Logic is not anymore exceptional, because, even in its meta-logical manifestations, it should not be taken to be the realm of the insubstantial or the uncontentious; logical theories (and their meta-logics) are substantial and contentious as any other scientific enterprise: therefore, they should be evaluated by means of an abductive methodology and judged "partly on their strength, simplicity, and elegance, partly on the fit between their consequences and what is independently known. […] Logic […] is no mere background framework but the very thing at issue" (423-24).[2]

Given this methodological framework, MLAM's main aim is that of convincing us that *necessitism*, the view that necessarily everything is necessarily something is the best metaphysics of modality. The endorsement of this *metaphysics*, however, is closely connected to (and probably is) an abductive endorsement of a *modal logic: higher-order necessitist S5*. Williamson shows that this logic, or better a specific intended model structure of it, is strong enough to give us the general structure of metaphysical modality, to give us substantial answers to a great numbers of modal metaphysical questions and it is better positioned than its opponents as far as strength, simplicity and fit are concerned. In showing this, Williamson achieves the impressive result of reconfiguring the entire debate on the philosophical foundations of modal logic and modal metaphysics. One way in which this reconfiguration happens is by abandoning (what Williamson believes is) the "badly confused" dichotomy possibilism/actualism and substitute it with the much clearer (at least for him) dichotomy contingentism/necessitism. Williamson's dissatisfaction with the possibilism/actualism dichotomy depends on the view that the only two plausible definitions of actual-

[1] Kaplan, D. 1994, "A Problem in Possible-World Semantics", in Sinnott-Armstrong, W., Raffmann, D. and Asher, N. *Modality, Morality and Belief: Essays in Honor of Ruth Barcan Marcus*, Cambridge: Cambridge University Press.
[2] See also Williamson, T. 2014 "Logic, Metalogic and Neutrality", *Erkenntnis*, 79, 211-31.

ism make this notion either viciously circular or trivially true (22-23). In the first case, the circularity depends on defining what is actual as what is true in the *actual* world, in the second case, the triviality depends on the logical behaviour of the actuality operator which could be inserted, *salva veritate*, in any non-modal context (at least in its standard interpretation); from this it follows that "actualism", the view that what is is what *actually* is, turns out to be a logical truth and possibilism, the negation of actualism, a logically falsehood. The entire possibilism/actualism debate is thus trivial and it should be substituted by the contingentism/necessitism debate which does not rely, at least explicitly, on any notion of actuality.

In this review, I will mainly be concerned with chapters of the book which make clear the sense in which a logic could be a metaphysics and justify, present and defend higher-order necessitist S5. These chapters could be taken as the "core" part of the book (at least in my opinion): action begins at Chapter 3 (from section 3.3) and continues through Chapters 5 (really from 5.5) and 6. Chapters 1-2 and 7-8 could be seen, respectively, as a preparation and an application of the core part.

Chapter 1 is devoted to philosophical preliminaries such as the definition of necessitism and of contingentism (as the negation of necessitism), to the temporal analogues of necessitism/contingentism, and to the critique of the actualism/possibilism distinction (which will be further elaborated in Ch. 7). Chapter 2 is devoted to the emergence of the Barcan Formula and its converse in the twentieth-century logical philosophy in authors such as Barcan Marcus, Carnap, and Prior: the chapter represents an extraordinary excursus on the pre-Kripkean developments of the semantic of modal logic. Chapter 7 is devoted to the possibility of "inter-theoretical communications" between necessitists and contingentists through the method that Williamson calls "of neutral equivalents". The communication between contingentists and necessitists is possible because contingentists and necessitists speak the very same language characterised by an unrestricted interpretation of the quantifiers and a metaphysical interpretation of modal operators. This proves, for Williamson, that, unlike the possibilism/actualism debate, the debate between contingentists and necessitists is non-verbal and substantial. Within the common language, there will be lots of sentences that are *neutral*, namely sentences over which the necessitist and the contingentist do not disagree, and there should be, at least in principle, a systematic way of "mapping" (where a "mapping" is not a translation) the non-neutral sentences into this neutral territory. Assume that there is a sentence D such that the necessitist accepts D and the contingentist rejects it; $(D)^{nec}$ is a mapping of D into this neutral territory (if there is such a mapping at all). Given that the necessitist accepts the equivalence between D and $(D)^{nec}$, the necessitist accepts $(D)^{nec}$, but given that $(D)^{nec}$ is, by hypothesis, neutral, also the contingentist can accepts $(D)^{nec}$. By accepting $(D)^{nec}$, the contingentist can thus extract some "cash value" from a sentence that she does not accept in its non-neutral formulation. It turns out that necessitism has more expressive power than contingentism, because, while everything that the contingentist can express can also be mapped into the neutral part of the language (i.e., where A is a sentence that the contingentist accepts, there is always a mapping $(A)^{con}$ that the necessitist can accept), the necessitist can express some theses that cannot be mapped into the neutral part of the language. Williamson thus concludes that "the necessitist can draw more distinctions than the contingentist can. Every distinction the contingentist can draw

can be drawn in neutral terms, so the necessitist can draw it too. The converse fail. The necessitist can draw distinctions the contingentist cannot, because they cannot be drawn in neutral terms." (364). Finally, Chapter 8 is a very rich and stimulating chapter devoted to the consequences of necessitism: among those worth noticing there is the view that the distinction between modal and non-modal dimension is elusive (and this saves the necessitist from the need to explain how the modal supervenes on the non-modal), the view that the incompatibility of necessitism with the truth-maker principle is a benefit of the view, rather than a cost and the view that necessitism is able (albeit with some difficulties) to account for radical contingency.

Let us now turn to the "core" part. Section 3.3 is a fundamental section of the book, because it makes clear in what sense a modal logic can be treated as a metaphysics of modality. What MLAM tries to establish is that the set of "sufficiently general" modal metaphysical truths has "the formal characteristics of a modal logic" (92). Of course the connection between a modal logic L and metaphysical modality cannot be direct. At least in the tradition of Kripkean model theory, the connection between a logic and a philosophical theory should be established by means of an *intended interpretation*. But an intended interpretation of a modal logic (one where, for example, to a certain predicate P is associated a specific predicate such as "being a tiger") is not sufficiently general. What Williamson is looking for are the general, structural principles of metaphysical modality and, at least in the propositional case, such principles are those that remain true under the permutation of their non-modal parts by holding fixed logical constants and modal operators in their metaphysical interpretation. This kind of permutation could be expressed, for Williamson, by means of quantification into sentence position. A formula like $p \to \quad p$ is thus a general structural principle of metaphysical modality (is "*metaphysically universal*", 93) if and only if the formula $\forall X (\quad X \to \quad X)$ is true. MLAM tries to show that there is a logic, *MU*, whose theorems are all and only the metaphysically universal truths and thus *MU* is "the uniquely correct logic of metaphysical modality". In particular, the formulas valid in an intended modal structure of *MU* are exactly the metaphysical universal ones and validity in such a model structure *is* metaphysical universality. The next step is to characterise the elements of such a model structure by means of a Boolean algebra of propositions; this move allows Williamson to show that *MU* is none else that S5, namely to show that all and only the theorems of S5 are metaphysically universal. An extension of the same techniques to the predicative case allows Williamson to define metaphysical universality for the first-order case (124). Here the fundamental argument is one where Williamson shows that a formula like $\exists y \, (x = y)$ is valid in an inhabited, pointed, model structure (a model structure with a domain and with the specification of an actual world) only if the Barcan Formula is valid in that model structure. The validity of the Barcan Formula (and, in the end, its metaphysical universality) is thus shown by Williamson to be simply a consequence of the validity and metaphysical universality of $\exists y \, (x = y)$. This is bad news for the contingentists, because this result shows that, if there is an intended inhabited structure for *MU*, then contingentism is false. The only option left for contingentists seems to be that of denying that there is an intended inhabited modal structure or having a totally instrumentalist approach to the model-theoretic apparatus.

Chapter 5 and 6 represents the higher-order extension of what Williamson has done in Chapter 3. According to him, a deeper comparison between contingentism and necessitism could be pursued in the contest of a higher-order modal logic where one can quantify both in predicate and sentence position. Within the most sensible system of higher-order modal logic, ML_p, a comprehension principle of the form $\exists X (\quad \forall x (Xx \leftrightarrow A))$ (whose role is that of guaranteeing a suitable array of properties) turns out to be metaphysically universal and thus valid in the intended modal structure (section 5.5, 230). Very briefly stated, the argumentative line of Chapter 6 is the following: necessitists are better positioned than contingentists to accept a robust interpretation of such a comprehension principle and to accept its consequences in a non-instrumental way. One of these consequences is *haecceitism*, the view that for every object, it is necessary that there is an haecceity of that object (where the haecceity of an object *o* is the property of being *o*). Another consequence is a second-order analogue of the necessary existence of individuals, namely the view that necessarily, every property is necessarily something (263). This conclusion is based on the assumption that necessary co-extensiveness of properties is the second-order analogue of first-order identity for objects. Actually, the conclusion is even stronger, because the comprehension principle implies necessary existence at *every* order except the first. The contingentists cannot accept neither haecceitism nor "any-order" necessary existence so they are forced to reformulate the comprehension principle. The problem is that weaker or non-modal formulations of the comprehension principle are, according to Williamson, unsatisfying: they would prevent "second-order logic from adequately serving the logical and mathematical purposes for which we need it" (288). On these abductive grounds, Williamson concludes that, if we want to use higher-order modal logic as a framework for our modal metaphysical inquiries, necessitist interpretations of ML_p should be preferred to its contingentist interpretations and thus necessitist higher-order S5 is our modal metaphysics. It has to be noticed, however, that the abductive grounds presented by Williamson do not seem to be so vast: the "mathematical purposes" served by a necessitist interpretation of ML_p are limited to some inferences within modal set theory, admittedly a quite remote and exotic field of mathematical practice (287).

The achievements of this book are impressive and they will surely change the way in which modal metaphysics will be done in the next decade. Before concluding my review, however, I would like to point out two problems. The first is about the elimination of the possibilism/actualism debate in favour of the contingentism/necessitism debate. Williamson claims that, on the operator account, "being actual had better be actually doing something harder than just being, otherwise the supposed dispute is silly" (23). I think that the critique is not entirely fair. As C. Menzel notices[3], the actualist should not be represented as someone looking for a robust conception of actuality, for a conception of actuality where actuality should be viewed as some "hard"-core feature of the world: rather, an actualist is *denying* the possibilist bifurcation of reality between *possibilia* and *actualia* (or between an heavyweight and a lightweight notion of existence) and she is simply denying the existence of *possibilia* (or the distinction be-

[3] Menzel, C., "In Defence of the Actualism/Possibilism Distinction", unpublished, https://www.academia.edu/20200724/In_Defense_of_the_Possibilism-Actualism_Distinction

tween two senses of existence). Assume that there is a philosopher who defends an analogous bifurcation between physical and "spiritual" things. A physicalist philosopher would be one who is simply denying the bifurcation by denying the existence of spiritual things. Sincerely, I do not see how this dispute (and by analogy the possibilism/actualism dispute) should be seen as "silly". So, I think that, after all, the dichotomy between possibilists and actualists will survive the publication of MLAM. The second problem is this. It is a recurrent theme of the book (specially section 3.6) that contingentists can have at most an instrumentalist attitude towards the model theory for modal logic. The problem for contingentists is that, if there is an intended model structure whose logic is sound and complete for metaphysical universality where $\exists y \, (x = y)$ is true, then, it is quite easy to prove that the Barcan Formula is true in that intended modal structure. Contingentists should therefore either deny that there is an intended model structure or hope that the model structure is only indirectly related to metaphysical universality. In Chapter 8, however, while defending the compatibility of necessitism with radical contingency, it turns out that necessitists too should have an instrumentalist attitude towards the intended modal structure: in particular, they should deny that the intended model structure is explanatorily prior to the contingent propositions from which it was constructed (410). So even necessitists should retain an instrumentalist attitude towards the model theory. Maybe, the instrumentalist attitude of the contingentist needs to be more radical than the one required to the necessitist, but given that the main aim of the book was to show how a logic could be literally taken to be a metaphysics, this instrumentalist residue towards the logic could be problematic for the necessitist as well.

University of Padua   VITTORIO MORATO

Vetter, Barbara, *Potentiality*.
Oxford: Oxford University Press, 2015, pp. ix+335.

Potentiality aims to provide a disposition-based account of modality, specifically metaphysical modality. Some interesting twists, however, separate it from other efforts in this direction. First, (Ch. 2 to 4) Vetter develops her preferred view about dispositions, then applies it to the treatment of modality (Ch. 5 to 7). These two parts are not at all disjointed. Firstly, her account of dispositions, and specifically their individuation, dictates the starting point in the treatment of modality: *possibility* rather than *counterfactuals*. Vetter does not believe dispositions to be primarily linked to stimulus/manifestation counterfactuals (Ch. 2). Refreshingly, she does not argue against a conditional analysis of dispositions through the usual interferences counterexamples (one may share the feeling that little to no progress can be achieved, on neither side, by pursuing this weary debate). Rather, as in Manley and Wasserman (2007, 2008),[1] Vetter discusses

---

[1] Manley, D. and Wasserman, R. 2007, "A Gradable Approach to Dispositions", *Philosophical Quarterly*, 57, 226, 68-75. And from the same authors, 2008, "On Linking Dispositions and Conditionals", *Mind*, 117, 465, 59-84. One can also consult the exchange between Manley, D. and Wasserman, R. 2011, "Dispositions, Conditionals, and

deeper, structural problems within the conditional analysis. In Ch. 3, using linguistic and lexicographic considerations as a jumping point, she develops a new account of dispositions, individuated by manifestations alone; their modal character is (to some approximation) that of possibility.

However, the *trait d'union* between the two parts of the book is the introduction of *potentialities*. Potentialities are a generalization and extension of our ordinary dispositional concepts, ranging from dispositions to powers, capacities, abilities and so forth. There is a number of reasons for their introduction. First of all (Ch. 3), potentialities serve as a vagueness- and context-insensitive metaphysical background for dispositional ascriptions; as such, they come in degrees of modal strength. Potentialities are also introduced to obtain suitable items for the definition of modal operators. The key-word here is "suitable", as Vetter introduces three constraints for a potentiality-based theory of possibility: *extensional correctness* (potentialities must deliver enough possibilities), *formal adequacy* (potentiality must have the right logical form and structure), and *semantic utility* (potentiality must reconstruct a good portion of modal semantics for natural languages as well). Ordinary dispositions, such as my irascibility, or the fragility of a glass, are not up to this task; thus, potentialities (in a number of varieties, as introduced in Chapter 4) function as suitable extensions thereof. Chapters 5 and 6 mostly deal with the second and third constraint; Chapter 7, discussing counterexamples, deals with extensional correctness. An appendix, finally, gives some formal details.

Vetter's goal is, primarily, a definition of the possibility operator in terms of potentialities. The definition has to capture the idea that for each possibility, there is a potentiality possessed by some (usually, but not always, concrete) object. Each possibility is thus a potentiality considered in abstraction from its bearer; the former, unlike the latter, is what Vetter calls *non-localized modality*. The definition, so to speak, tells us where to look to find modality in the world: in potentialities possessed by objects. Vetter urges us to think of the localization of possibility operated by potentialities as analogous to the localization of necessity operated by Finean essences.[2]

Potentiality presents a great deal of interesting issues that cannot possibly be covered in this review. There is, however, a big problem for any account of possibility in terms of dispositions, that *Potentiality* brings to the forefront, perhaps for the first time; its resolution occupies a significant portion of the book, and the strategy adopted is, I believe, especially worth discussing. This problem has both formal and informal components, and comes in the form of a gap to be bridged between dispositions and possibilities. The driving idea behind a treatment of possibility through dispositions is a biconditional schema such as:

(D) It is possible that x is P iff x has a disposition to be P.

If the goal is a definition of possibility, **(D)** can only go so far. Not only it is open to many counterexamples (what happens if x does not exist?); a more general problem is that possibility is standardly formalized as a sentential operator,

Counterexamples", *Mind*, 120, 480, 1191-1227, and Vetter, B. 2011a, "On Linking Dispositions and Which Conditionals?", *Mind*, 120, 480, 1173-1189.

[2] However, she discusses (164 ff) the relation between potentialities and Finean essences to the conclusion, that, unlike possibility and necessity, they are not duals. What the precise relation between them is, is left unsaid.

whereas dispositions are naturally linked to predicate modifiers—as in **(D)**. In short, it is not straightforward how to generalize **(D)** into an explicit definition of the diamond operator.

This is where Vetter's strategy of introducing potentialities kicks in; it consists, as she repeatedly points out (e.g., 141) in taking "the path of least formal resistance"; viz., she will help herself with the simplest formal devices to close the logical gap in **(D)**, and only in a second moment she will worry about the metaphysics backing such a formalization. The formal work is indeed quite simple: Vetter allows sentences of arbitrary logical complexity into the scope of a potentiality "POT" predicate modifier, then uses a lambda abstractor to obtain predicates to put in it. This is why most of the potentialities described display the peculiar form "the potentiality to be such that ϕ", or "the potentiality for ϕ". The quantity of potentialities allowed by this strategy is astounding, but Vetter spends considerable energy (throughout Chapters 4 and 5, primarily 148ff) to assure the reader that, from a metaphysical standpoint, they "make sense". E.g., how can an arbitrary *a* have a potentiality to be such that (distinct) *b* is F? That this is a bogus potentiality is suggested by the fact that in the suggested formalization POT[λx.Fb](a) there is no free occurrence of x in the scope of the lambda abstractor. But maybe there is some sense to be made of some *a* having a potentiality for *b* to F; such a potentiality will more likely be extrinsic (for its possession by *a* entirely depends on how *b* is)—and will presumably depend on some further potentiality that *a* and *b* jointly possess. Additionally, in virtue of *a* having a potentiality for *b* to be F, *a* also has a potentiality for something to be F (viz., POT[λx.$\exists$xFx](a))—thus, quantifiers inside the scope of POT are fine as well. Vetter offers some examples, although—as we move farther and farther away from ordinary dispositional ascriptions—they are not as intuitive as perhaps the reader would have hoped. Through potentialities the expressive reach of ordinary dispositions is considerably extended, helping dispositions meet both the *formal adequacy* and the *extensional correctness* constraint. The final step is the admission of iterated potentialities, viz., potentialities whose manifestations are, or involve, the possession of other potentialities. There is no general pattern in this "potentiality-finding" operation (Vetter clearly favors piecemeal approaches to catch-all solutions, e.g., in dealing to counterexamples in Chapter 7). However, she argues negatively (155-157) that any restriction of acceptable potentialities based on their logical form would be highly arbitrary. Eventually, a schematic definition of possibility (197) is formulated:

(D₁) $\Diamond\ \phi =_{df} \exists$xPOT*[$\phi$](x),

where POT* is the sentential operator for iterated potentiality; the *definiens* of **(D₁)** must be read as "something has an iterated potentiality for it to be the case that ϕ". The existential quantification expresses the localization of possibility: modality is tracked down to at least one specific object, possessing a specific potentiality.

The ontology of potentiality is thus generous. Vetter rejects non-trivial necessary conditions for object(s) having joint potentialities (117ff). Spatiotemporal proximity or causal interaction are no matter: one can have a joint potentiality to sing a duet with their best friend, or with Queen Elizabeth (however, *which* potentialities an arbitrary collection of objects can have is presumably subject to some restriction; 175). Furthermore, Vetter frequently suggests that many object(s) that we would say do not possess a given potentiality, actually *do* possess it to a very low degree—perhaps too low to warrant a dispositional ascription.

One might believe this be a drawback. Not only for parsimony considerations, but also because it seems nothing is left of the reassuringly "naturalistic" ontology of dispositions (23). To be sure, this motley crew of potentialities is not entirely left unruly. We cannot stipulate *a priori* which potentialities are instantiated. First (202), because in most cases, this is the job of empirical enquiry.[3] Secondly, aside from tautological potentialities—which are trivially possessed across the board—Vetter never formulates sufficient conditions for some object(s) possessing potentialities; however (114), some restriction is clearly in place, since it is not the case that everything has every potentiality with everything else.

Unsurprisingly, many, perhaps most, of the potentialities needed to treat modality will reveal themselves as non-fundamental potentialities. Vetter embraces this fact: "we are not formalizing the potentialities that are metaphysically basic. We are formalizing any potentiality, no matter how derivative." (142). This is, to my eyes, one of the greatest strength of her account, but it also draws attention to what one can and cannot accomplish through it. One can see its strength: as she very briefly points out (142), predicate logic does not care about the metaphysical status of properties expressed in it. Something similar, I would say, holds in modal logic as well; e.g. **(D₁)** is a schema in which ϕ can be replaced by any well-formed formula, so one can reasonably expect all kinds of gerrymandered possibilities. The moral of the story appears to be that if the goal is to close the formal and informal gap in **(D)**, and obtain an extensionally correct definitional biconditional, one needs all kinds of derivative potentialities as well (195 offers the what could be a slogan: "every potentiality counts"). It is not something that Vetter lingers upon, but *Potentiality* appears to be making the interesting point that if the *desideratum* is a definition of possibility, one should not have a fixation with fundamentality (e.g., 195-98).

Relatedly, one way to appreciate strengths and limitations of the approach is by considering other non-definitional strategies for a disposition-based account of modality. Someone interested in what is fundamental about modality, might decide to opt for a grounding or truthmaking claim; for instance, the claim that modal statements have dispositional truthmakers. But here is the catch: if the quest of truthmakers is to unveil some dispositional "deep story" about modality, the sub-propositional logical form of the truthbearer is not going to provide any insight on the matter. For one may suspect that, say, a disposition for P might be ultimately responsible for the truth that ◊ϕ even without any logical correlation between P and ϕ. Unfortunately, there is not much else that a mere examination of "ϕ" could reveal about its truthmakers (some would claim that there is *nothing* the examination of "ϕ" could reveal about its truthmakers). E.g., what ultimately makes it possible that the Sun rises in the West? It might be all manners of dispositions. The truthmaking claim does not offer any clue; one needs to develop a (presumably *a posteriori*) epistemology of grounding/truthmaking to find out. This holds for every account that abandons a one-to-one correspondence between possibilities and potentialities (and thus, any hope of a *definition)*, in order to focus on the fundamental potentialities. Surely such an account would not be as bothered by the logical form of potentialities (212) if it were just to claim that for each possibility, there is *some* potentiality or another that accounts for it. But then again, *which* potentiality?

---

[3] The author's epistemology of potentialities/possibilities is a largely *a posteriori* matter—something that Vetter considers "an advantage rather than a drawback" (268).

This puts Vetter's theory at an advantage when it comes to *informativeness*. For Vetter's definition can be applied, and thus affords somewhat of an "algorithm" that pairs every possibility with a potentiality. Mind you: through **(D₁)** one only knows that something has an iterated potentiality to ɸ conditionally on the knowledge that ɸ is possible (and vice versa: as Vetter discusses in relation with the *extensional correctness* constraint, it is a tricky matter to know whether ɸ is possible before applying the definition). This algorithm, and thus, Vetter's definition, is not a device to find potentialities *a priori*, nor to discover a dispositional "deep story" about possibilities; rather, it is a useful biconditional link between *specific* possibilities and *specific* potentialities.

Assuredly, given this "path of least formal resistance", the switch from possibilities to potentialities may appear to be a pointless manoeuvre: they are, after all, two logically isomorphic operators. According to Vetter, however "this general notion of potentiality has not been formulated directly, or simply stipulated out of nowhere; it has taken me four chapters to reach the generalized notion" (194-95). The reasoning seems to be that the shift from possibility to potentiality, admittedly pointless if merely a logical stipulation, acquires significance insofar as potentiality is a metaphysical generalization of more ordinary dispositions and capacities. Whether or not this is sufficient to blunt the charge, is left to be seen. Another limitation of Vetter's theory appears to be that—exactly because of the overabundance of derivative potentialities—it may be *per se* incapable of completing the desired localization of possibility. Consider the possibility that Lorenzo is a musician; it appears to be located in Lorenzo's potentiality to be a musician (viz., for Lorenzo to be a musician). Yet there is a staggering number of other entities having a potentiality for Lorenzo to be a musician, e.g., all my possible teachers, or all the devices I can use to learn. There is nothing wrong with multiple "witnesses" to this possibility (as there can be more than one variable value turning an existential quantification true, 201). This however means that the modality in "possibly, Lorenzo is a musician" has been equally localized in entities all across the world. I doubt that this is what the author intended. One may have the idea that ultimately, the possibility of Lorenzo being a musician is localized in Lorenzo's potentiality to be a musician; perhaps because only in Lorenzo the potentiality for Lorenzo to be a musician is intrinsic (some characterizations of *localized modality* appear to have something to do with intrinsicness, e.g. 103). For Vetter, Lorenzo's intrinsic potentiality is preferable to the extent that is needed to ground all the others. But, in general, the selection of basic potentialities is not something that transpires from **(D₁)**, nor from Vetter's logical framework for the treatment of possibilities.

Overall, what might be a weakness of this "path of least formal resistance" is that metaphysics is clearly taking the backseat, and is left to deal with the results of formalization, whichever those might be. This is a bit odd, considering that the reasons for this whole enterprise come mainly from metaphysics: one develops a disposition-based account of modality, even if possible world semantics is an extremely powerful tool, in order to have a more desirable metaphysics of modality (as the author maintains in Ch. 1).[4] Yet many potentialities (especially the iterated ones, or those with logically complex manifestations) are metaphysically suspicious. This complaint could take many forms: e.g., is it harm-

---

[4] And earlier, in Vetter, B. 2011b, "Recent Work: Modality without Possible Worlds", *Analysis*, 71, 4, 742-54.

less to quantify over and refer to non-existents in the scope of POT, as Vetter does throughout the book? POT, like ◊, is not factive in the intended interpretation—so one is not straightforwardly committed to such shadowy beings. However, one may think that they should excluded from having any role in the individuation of potentialities, even if manifestations are described by quantifying over, or referring to them. Vetter does not spell out what the metaphysical nature of manifestations is; nor in what sense they "individuate" potentialities. So it is not easy to formulate an argument; but worries still loom.

These doubts do not detract from the value of *Potentiality*. As innovative and detailed as it is, it is bound to raise new questions and breathe new life into old ones. For a crucial thing to notice about this book, is that it is a book. As of now, no other dispositional account of modality is available, as thorough and extensive as this one. Thus, in a way, *Potentiality* wears one of its merits up its sleeve. Vetter does not claim hers to be the only one viable dispositional account of modality, nor the best; she employs little to no time discussing alternatives, nor comparing her account to them: she is almost entirely interested in developing her option. The book shines so much in that department, that it is easy to see it setting somewhat of a bar in the field, that competing accounts will have to measure up to.[5]

Scuola Normale Superiore of Pisa LORENZO AZZANO