Table of Contents

Essays

Focus

Essays on Barbara Vetter's *Potentiality*

philinq VIII, 1-2020
ISSN (print) 2281-8618-ETS

Past Present

Essays

Spinoza: reasoned indifference as an introduction to adaptation in unusual circumstances

Abraham Mounitz

Abstract: The study offers a practical model of Spinoza's behavioral teachings for empowering a person's ability to cope in stressful situations. The concept of different and indifferent exerts significant implications on everyday lives of a person. This model offers philosophical tool for the intellectual control of emotions that weaken a person's power to act. The first part offers Spinoza's metaphysical basis, focusing on the two titular concepts that represent humans and nature. Next is Spinoza's philosophical method of guiding people toward conduct that is associated with and derived from nature's reasoning as values that ameliorate everyday conduct. The practical layer of this study offers a basic model, a philosophical anchor, which can be used for the formulation of empirical research questionnaires on various topic associated with an individual's adaptation to a challenging emotional environment and all it entails (feelings, ability to function etc.). The study also present a sample questionnaire formulated according to the Spinozist model. The study's final part presents several interviews conducted by the author in the model's spirit as an outline for future empirical studies and for the formulation of curricula designed in the spirit of Spinoza's behavioral philosophy.

Keywords: adaptation; indifferent; metaphysic; nature; passive-active.

Abbreviations: Part in the Ethics: E. (1-5); Appendix: app.; Axiom: ax.; Corollary: c.; Definition: d.; Definition of affects – da. ; Preface: pref.; Proof: pro. ; Proposition: p.; Scholium: s.

1. *Introduction*

The present study seeks to offer an applied model of Spinoza's behavioral teachings. This model has been used for several years in what is called "philosophical therapy" (see Appendix 2) and as such might serve as a practical basis for empowering a person's ability to cope in unfamiliar situations, and in stressful or embarrassing situations in particular. More specifically, I seek to offer an infrastructure for the development of further studies followed by the formulation of applied teaching plans pertaining to adaptations to the aforementioned situations.

philinq VIII, 1-2020, pp. 9-36
ISSN (print) 2281-8618-ETS DOI: 10.4454/philinq.v8i1.233
Submitted: October 2018
Accepted: July 2019

The central motive of Spinoza's *Ethics* is the empowerment of a person's abilities, and - by so doing – the promotion of this person's happiness. It is in this spirit that the relation between the concepts of different and indifferent exerts significant implications on the everyday lives of a person as an individual existing within an infinite totality. The methodological basis offered below is meant for guiding humans toward better functioning in cases of emotional ambiguity, nebulosity or a sense of helplessness. In other words, it may be seen as seeking to offer philosophical tools for the intellectual control of emotions that weaken, and occasionally even extinguish, a person's power to act.

Nature, being a totality containing all its individual parts, has no defined objectives, goals, or purposes. It acts according to the rules of its nature (the literal rules of nature) and as such may be referred to as an "active is". In other words, nature's action expresses its substance (what it is as a whole). Seen objectively, therefore, the actions of nature are not carried out for any particular purpose, and it is thus free as a whole and unencumbered by the desires or whims of any factor in the universe as a whole or of a human in particular. Indeed, one of the most powerful supports for Spinoza's identification of God with nature is the distinction between these two concepts. As infinity, nature is in general indifferent and indifferent in particular to our wishes and destiny, an indifference which necessarily follows from nature's freedom of action, from acting without a defined purpose, without coercion and with no meaning or any distinction between good and evil. Spinoza explicitly uses the word "indifferent" in relation to nature-God (E.1: 33, s.2). We, as human beings seeking to continue in existence, make a subjective and utilitarian distinction between good and evil. Spinoza teaches us to mimic nature's characteristic conduct insofar as is possible and to strive to act in accordance with our inner reasoned nature, which forms part of nature in general and which can understand and internalize the regularity of nature given that we are not a realm separate from nature, and not a lone island within it. This understanding imbues us with a sense of freedom in and of itself since it allows us to purify ourselves and dispose of those emotions that drive us toward passivity. This understanding in itself further transforms us into more active subjects acting under the guidance of a wisdom that matches our specific nature as an inseparable part of nature in general.

2. *From metaphysics to values*

Spinoza's method is clearly practical since its stated goal is to lead the person, tossed and turned uncontrollably in the tumultuous sea of events, to a safe haven of happiness. As such, and against the background of the connection

between metaphysics, ethics and psychology, it has recognized reason as an ability derived from nature in general since it accords with nature's reasoning and it is nature that makes it possible for reason to follow nature's rules and make use of them in advancing toward that safe haven of happiness (Ben-Shelomo 2012: 155).

According to Spinoza's method, the metaphysical relation between finite and infinite, as well as between part and whole, constitutes a descriptive model for the relation between human beings and nature as a whole. It therefore follows that – from the present study's perspective – the issue of the relation between different (human beings as individuals and a defined part of the natural totality) and indifferent (nature as a whole) enfolds the method's objectives: the indifferent represents the metaphysical component – the totality, while the different represents that the psychophysical and value laden part of the human individual within the method.

It is well known that Spinoza's teachings as a whole (metaphysics, psychology and ethics) are intended as a form of applied idealism rather than a theory meant to remain in an ivory tower bookshelf. The explanatory model presented here is nothing more than the depiction of a practical outline offered by Spinoza for everyday human conduct as is necessitated by the basic logic of human nature common to human beings as whole. "Spinoza's *Ethics* supports a functional connection between the metaphysical discussion concerning the deterministic structure of nature that was discussed in Part 1 and the remaining parts of the method that discuss behavioral and educational aspects" (Mounitz & Berenson 2016).

Insofar as Spinoza and the proposed model is concerned, nature – as a ceaselessly and endlessly active whole – is a rational ideal that is to be internalized and imitated to the best of our ability. To this end, and even at this early point in the discussion, I wish to present an extreme example where Spinoza determines that pity is a negative emotion: "In the man who lives by the guidance of reason, pity in itself is bad and disadvantageous". (E. 4: 50)

Pity gives rise to sadness, and sadness reduces a person's effectiveness in action. Let us consider, for example, a medic, an ambulance driver, a surgeon or an everyman encountering an unpleasant situation (an accident or some other medical event) where she or he is capable and even morally obliged to assist, and pitying the injured person or person lying before her or him. Indeed, I do not believe it is necessary to engage in an extended discussion of the implications of such pity. Similar examples may apply to anger, hatred, jealousy, competitiveness, belittlement, and other kinds of stressful situations.

3. *Different and indifferent*

The behavioral model offered herein, has already been tested and proved effective (See examples in Appendix 2) and relies on Spinoza's text, such as Letter 50, which states that any definition expresses a negation and thus constitutes a limitation.

> […], since substance is by nature prior to its affections [all things], (Pr1), disregarding therefore, its affections and considering substance in itself, […] it cannot be conceived as distinguishable from another substance (E.1: 5, pro.).

In the physical sense, the definition of any object or form, such as a triangle, a plot of land, a house, a lake, etc. distinguishes it from any other thing: differentiated and different = different from anything that it is not. The physical aspect of non-indifference discussed in the metaphysical part of the *Ethics* is described by Bennet (1996) as a delimitation of proximity and thickness. The definition of attached, proximal, or joined bodies is thus contingent on the absence of motion or stagnation, and with any changes requiring their redefinition.

On the other hand, the definition of nature as a whole (substance) proceeds from negation: "Absolutely infinite substance is indivisible" (E.1: 13). In other words, we are concerned with the absence of the ability to compare nature to anything else given that nothing exists outside it and nothing comparable to it exists in any case, which is why it is "Absolutely infinite", i.e. undefined physically or in sense of its spatial expansion. These data lead to the logical derivation of two opposing traits: different and indifferent. God – an object defined negatively – is in fact undefined neither in space or time since it does not relate and is not relatable. Spinoza identifies God with nature (*Deus sive natura*), and so his use of "God" actually means the infinite totality of nature. It should also be noted that the term 'indifferent' formally subsumes the term 'infinite' - that is to say indistinguishable by size or limits, and by so doing completes the descriptive picture of the unlimited spatial definition within Spinoza's metaphysics. Put differently, it may be argued that God's laws are nature's laws rather than the laws of a personality that manages the universe from the outside. We are thus concerned with a necessary order – one aspect of which also includes a person's emotions, thoughts, decisions, and actions as part of her or his reality and as a result of humans singular and unexchangeable place as an individual part of a totality (Rutherford 2010: 143-167; also see Curley 2013).

The text and the background of an interdisciplinary connection between metaphysics, psychology and the method's morality lead us to infer that God's indifference is not merely the lack of a formal definition (due to the impossibil-

ity of forming a reference point) but is rather an inward facing projection relating to anything derived from it and exists within it: "Nature has no fixed goal and that all final causes are but figments of the human imagination". (E.1, app.).

In the pref. to P.3 Spinoza declares that "I shall consider human actions and appetites just as if it were an investigation into lines, planes, or bodies". Spinoza uses the term "indifference" to describe nature's behavior in the mental sense (E. 1: 33, s.2). The Latin dictionary definition of *indifferensentis* states that it is "neither good nor bad", "not fussy or particular, indifferent" (Oxford Latin Dictionary 2012). A person, like all methods derived from and defined within the one and only totality, cannot ignore what seems to be good, bad, beneficial or harmful according to her or his view of the benefit or harm reflected unto her or him in the world of phenomena. It is not merely that a human, like any other body, is defined and distinguished spatially from anything external to it, but also that humans require comparisons, and as such are defining, separating, joining and comparing creatures that are not indifferent to the objects of these activities, i.e. that care and do not treat things with the same cold and rational indifference in which nature conducts itself with respect to humans.

This asymmetry between person and whole we have witnessed is more than a mere metaphysical problem derived from a differentiation of means (as part of the multitude of phenomena) within the substance (nature) in Spinoza's *Ethics*. With respect to the present perspective, Spinoza offers a certain correction, a kind of reasoned remedy, concerned with education in general and the reader's self-education in particular. According to Zourabichvili (2002), education is a continuous transformation from childhood to adulthood and continuous refinement to achieve the desired goal in the *Ethics*.

The study proceeds the way to adaptation through reason. This key objective paves the way toward grounding the article's central arguments, viz. (1) nature's reasoning is indifferent to values from both a metaphysical and a scientific perspective; (2) scientific examination is indifferent to the illusion of discrete time; (3) any reasoned view is bound to be deterministic and as such supports rational indifference as does determinism itself; (4) nature's indifferent treatment of events - as a model to be emulated – plays a key part in the life experience of the reasoned human being in unconventional situations.

This would also be the place to emphasize the role and implications of geometric formalism in Spinoza's method. It is well known that the *Ethics*' geometric structure reconciles the requirement for consistency with the method's rational content (Steenbakkers 2009; Barbaras 2007). Moreover, it is important to note the indifference projected on the content by Euclidean geometry, i.e. the formal frameworks indifference to its content. According to Delassus (2018), the geometric structure of the *Ethics* as an explanatory tool for man's

behavior, desires and actions expresses a frame devoid of emotions for the reasoned explanation of what is outside us, and thus Spinoza treats man as an inseparable part of nature (6, 7). In this spirit, I believe that the geometrical framework as an explanatory tool of the *Ethics* carries with it a message of an individuated reference as a starting point for understanding the world (nature) indifferent to our feelings. O'Donnell (2018) states that part of the power of Spinoza's thinking stems from the unconsciousness and alienation we are required to undergo as part of our understanding of our situation (826).

It is through these measures that the study intends to extract a conclusion suggesting that the method's indifferent perspective plays a practical part in decision making and in individual functioning under stressful conditions where a person is required to be of use to others. This reasoned indifference is thus a virtue that allows a balanced view of reality. It is through the lens of a scientific and temporal indifference that it becomes possible to perceive things objectively and to view oneself as part of a complete and determinist worldview. In other words, the meaning of the presence of nature as a whole as a guiding principle is indifferent to its contents is its being an object of striving, understanding, and even imitation and internalization. This presence of a totality that is indifferent to our destiny and our values thus guides the reasonable person toward an embedding of this guiding ideal in her or his consciousness. The results of such conduct are known. For example, such a person could peacefully bear the vicissitudes of faith since "…in so far as we are intelligent beings, we cannot desire anything save that which is necessary, nor yield absolute acquiescence to anything, save to that which is true: wherefore, in so far as we have a right understanding of these things, the endeavour of the better part of ourselves is in harmony with the [indifferent] order of nature as a whole…" (E.4: 32, app.). Spinoza's principles thus offer a practical way of rational conduct in everyday life guided by the metaphysical principles of nature.

It should be emphasized that earlier field studies conducted on the basis of Spinoza's principles proved the existence of a basic common denominator among all humans with respect to the cognitive skills involved in intellectual attention, self-control, and implementation ability. This is why every human has the potential for self-preservation as well as the ability to internalize metaphysical principles insofar as they are understood to have been promoting the objective of self-preservation – i.e. insofar as they are capable of being cognitively applied if necessary (Mounitz & Berenson 2016). Delassus (2014) refers to the *Ethics* as a strategy for dealing with situations of illness.

The model proposed in this study has been used successfully for several years to deal with problems of tension between partners, anger, sadness, jealousy, and desire for revenge, and can therefore serve as a basis for empirical research in

medicine and in situations in which these conditions exist (See Appendix 2).

The term 'adaptation', which has featured prominently in behavioral and educational research in the past few decades (Vaillant 1993) is a psychological representation of what Spinoza refers to in the *Ethics* that was not written 350 years ago as a result of empirical research but rather as the result of observing human behavior as well as a result of Spinoza's self-reflection. I believe that the present study's proposed model could in turn serve as a baseline for an educational method aiming for a reasoned kind of this so-called 'adaptation', i.e. for an adaptation to reasoned balance in an emotionally evocative environment. By so doing, we shall be continuing the work we have begun several years ago, viz. taking philosophy down from 'ivory tower' arguments to the 'factory floor' of the empirical reality of everyday life in the spirit of the arguments in Golomb (2015).

4. *Metaphysical division: the reasoned indifference of nature-god*

Spinoza's linking of the metaphysical and psychological-behavioral parts of his philosophy in the *Ethics* is functional in nature. The geometrical method that guides his metaphysical discussion of natural laws is equally employed for the arguments he posits in the psychological part of the *Ethics* (Delassus 2014). Naturalist regularity can thus be understood and applied in everyday life and it is also the factor that determines our position as individuals within this metaphysical totality (LeBuffe 2010: 28-29). According to Spinoza, we cannot know what we are – our limits, our powers, or our "good" – unless we have an understanding of the whole (O'Donnell 2018).

This ability to understand naturalist regularity is common to all human beings and is instilled through the rationales of Euclidean geometry, whose basic logic cannot be opposed by any sentient being.

Some (mistakenly) ascribe God with some dimension of length, breadth and depth that is limited by the surrounding environment. Stating such things about the infinite 'is' is, however, entirely false (E.1: 15, s.); or, in Spinoza's own words: "God acts solely from the laws of his own nature, constrained by none" (E.1: 17).

It therefore follows that (a) there is no external or internal cause that motivates God to action beyond the perfection of his own nature; (b) God is necessarily a free reason that is present and active and that "nothing in nature is contingent, but all things are from the necessity of the divine nature determined to exist and act in the definite way". (E.1: 29); (c) a central statement the indifference principle relies on is E.1: 32, which reads: "Will cannot be called free cause, but only a necessary cause".

In other words, will represent a conscious inclination directed by a sense of judgement that distinguishes between good and evil and always cares (compares, measures and is not indifferent). Spinoza proceeds to argue that "it follows, secondly, that will and intellect, bear the same relationship to God's nature…" and that "things could not have been produced by God in any other way or in any other order than is the case" (E.1: 32, c. 2 and E.1: 33).

Put differently, Spinoza is arguing that nature as a whole necessarily reflects perfection and order, that there is no good and no evil and that there is no volitional purpose to natural actions. Nothing can occur other than what exists and thus it must not be determined that reality is the result of any kind of natural volition.

God or nature as a substance this is the first and only free cause for both the existence as well as the essence of all things including human beings. Humans, like any other things in nature are God's derivative (*modus*) that is necessarily present in God-nature (Della Rocca 2008: 70). The ascription of any kind of purpose to divine actions negates God's perfection since it follows that God is yearning for something that he lacks. The projection of human will on the nature of God-nature is thus nothing but a safe haven for unknowability (E.1: 32, c. 2 and E.1: 33). The thing which guides the human non-indifferent view of things is what they perceive as useful or what causes the greatest pleasure. It therefore follows that humans have determined the concepts of good and evil, order and chaos, hot and cold, ugliness and beauty, early and late, etc. because they perceived themselves as possessing a sense of choice. It is as beings possessing a sense of choice, therefore, that they projected their non-indifferent view on God-nature.

5. *A clarifying eemark on fatalism*

Indifference, insofar as the term is used by Spinoza, is presently discussed in its positive sense. We are not concerned with disinterest, fatalism or unaware apathy – not at all. Spinoza avoids these senses of the concept explicitly and even instructs us to reject anything that might interrupt our enjoyment of "reasoned life in the manner we consider safest" (E.4, ap. 8).

Positive indifference is a clearly rational product arising from nature as a substance and from which a reasoned person may drink deeply in the service or her or his personal salvation. Indeed, in E.2: 44, Spinoza himself argues that "it is not in the nature of reason to regard things as contingent, but as necessary".

Spinoza attacks apathy and disinterest and considers both as a kind of helplessness arising from ignorance and stupidity. If reason is the key to self-existence and to the maximization of vital activity, then apathy is its antithesis

and expresses a type of conduct that goes against human nature and as such is referred to as a disease that goes against self-preservation (Green 2016).

Spinoza thus employs this starting point to convey a clear educational message. If nature's conduct serves no purpose, and if it is the reason for all of the phenomena it (and the world) contains, then it follows that anything that happens is necessarily caused by the reality of nature being its own reason without aspiring to any particular purpose. Any view of things as coincidental arises from images of good and evil with respect to the phenomenon, and – as such – is wrong or partial and certainly never fatal. Spinoza rejects fatalism and compares the fatalist to a donkey who is equally hungry and thirsty and who shall die when it reaches green pastures and a spring on account of a lack of ability to decide whether to begin by eating or by drinking (E.2: 49, s.).

Fatalism may lead to suicide (Green 2016) and is nothing but mental fatigue and total submission to external causes. Similarly, compassion expresses an emotion of sadness to the extent of helplessness. On the other hand, reasoned indifference is a clearly active act that removes negative emotional traces and actually evokes action. Medical professionals or any person who is not enslaved to negative emotions will function with greater efficiency when she or he is required to cope with a concrete case brought before her or him due to a deterministic view of the situation. Indifference is thus a reasoned philosophical tool which manifests what psychologists refer to as "adaptation", as being resigned to the human environment and as indicating the actor's mental maturity (Vaillant 1993).

Spinoza argues that our cooperation with nature increases with the perfection of the acts we engage in since this perfection allows us a greater familiarity with nature which is a triumph of our minds (E.2: 49, s.). He also instructs us on how to behave in the severe eventualities we encounter in our everyday lives: "What attitude we should adopt regarding fortune, or the things that are not in our power, that is, the things that do not follow from our nature; namely, to expect and to endure with patience both faces of fortune. For all things follow from God's eternal decree by the same necessity as it follows from the essence of triangle that its three angles are equal to two right angles" (E.2: 49, s.).

In this case, the expression "to endure with patience" refers to a form of emotional expression but also to a clear product of reason, the conscious internalization of the metaphysical context of irregular events we encounter in our everyday lives. We are obviously not concerned with the product of psychotherapy, psychiatric care or the use of chemicals, alcohol, drugs and medications. Spinoza is thus referring to reasoned activism as a tool for "endur[ing] with patience", and the social benefit thus derived is a side effect of the utility derived by an active person who shifts from passivity (or a low ability to act)

to activity. Kisner (2008) suggests that Spinoza does not view passions and endeavors as negative since they arise from the law of self-preservation which is informed passion (for existence) in and of itself. An endeavor may thus be seen as promoting the human aspiration for self-preservation – for existence – insofar as it is reasoned. Kisner stresses that Spinoza cannot accept the ethicality of apathy and rejects the interpretive tone which associates Spinozan thought with Stoicism that argues that impulses may be completely neutralized through reasoned control. Kisner argues that Spinoza presents impulses as playing a central role in human existence. According to Spinozan thought, therefore, humans may use reason to maximize their awareness of their impulses' existence and to consequently direct these impulses to the functional framework of "intelligent conduct" (759-783).

Reasoned indifference is thus a conscious action and the result of an extremely high degree of self-awareness: "We are active when something takes place, in us or externally to us, of which we are an adequate cause" (P.3, d.2). In other words, if we are sad as a result of an external impression that affected us, we exist in a state of passivity since this sadness is not necessitated by the fact of our existence but rather by an external influence which exerts a negative effect on our nature that seeks continues existence. Non-indifference is what creates the partition, the border, and the wall between us and the world outside us.

The emotions which explain joy and – conversely – which explain sadness are those which shape the changes in our minds. The mind may accept great changes and move at times to great perfection (as in joy) and at times to a lesser perfection than that which preceded it (in sadness) to the point of helplessness. In E.3: 13, Spinoza states that the consciousness can seize and move away anything that reduces or inhibits a person's force of action. In other words, the mind refuses to picture that which reduces or inhibits a person's bodily and mental strength. Spinoza later defines compassion: "Pity we define as 'pain arising from another's hurt'" (E.4: 22, s.).

Further along, Spinoza expands his discussion of this emotion and states that "pity is pain accompanied by the idea of ill that has happened to another whom we think of as like ourselves" (E.4, da. 18).

He notes that compassion not only arises in us with respect to people we love, but also in relation to people we have never come into contact with. In E.4: 50, c. and s. Spinoza emphasizes faults in compassion and language: "in the man who lives by the guidance of reason, pity is in itself bad and disadvantageous".

He also explains his reasoning:

> Pity is a pain and therefore in itself is bad. Now, the good that follows from it (that we endeavor to free from distress one whom we pity) we desire to do solely from the

dictates of reason, and it is only from the dictates of reason that we desire to do something that we certainly know to be good. So, in the man who lives by the guidance of reason, pity is itself bad and disadvantageous

and in Corollary:

hence it follows that the man who lives by the dictates of reason endeavors, as far as he can, not to be touched by pity.

As noted, we are not concerned with fatalism, but rather the opposite, since Spinoza is in favor of mutual aid, but not as a result of pity or superstition but as a result of the governance of reasons: "For Spinoza, when I act out of pity, I am striving to ease my own suffering which involved in that very feeling of pity. In this way, Spinoza would characterize an apparent case of altruism as one that does not involve altruism at all" (Della Roca 1996: 232).

Spinoza instructs us on the manner of avoiding the effect of this emotion: via the reasoned indifference referred to as "the guidance of reason" – or, in other words, by copying nature's metaphysical indifference. As stated by Spinoza itself:

He who rightly knows that all things follow from the necessity of the divine nature and happen in accordance with the eternal laws and rules of Nature will surely find nothing deserving of hatred, derision, or contempt, not will he pity anyone. Rather, as far as the virtue of man extends, he will endeavor to do well, as the saying goes, and be glad. Furthermore, he who is easily touched by the emotion of pity and moved by another's distress or tears often does something which he later regrets, both because from emotion we do nothing that we certainly know to be good and because we are easily deceived by false tears. Now I emphasize that I am here speaking of the man who lives by the guidance of reason. For he who is moved neither by reason nor by pity to render help to others is rightly called inhuman. For (E.3: 27) he seems to be unlike man (E.4: 52, s.).

The comment's end reinforces Spinoza's approach in opposing indifference and apathy toward others as discussed above. He only rejects unreasonable acts since they are driven by sadness and since they lead to sadness. His next sentence proceeds to immediately defining favor, which is the opposite of compassion since it can accord with reason, In Spinoza's terms: "Approbation (favor) is not opposed to reason; it can agree with reason and arise from it" (E.4: 51).

Zahavi (2015) argues that empathy is not sharing or participating but rather a basic kind of sensitivity to and understanding of the other with a view to bettering others' existence. It is thus not a form of social behavior but rather a kind of precondition for sharing. Zahavi thus views empathy as an intel-

lectual ability to assess another person's condition at a glance. This intuition, in turn, also acts as an initial starting point for interacting with others for their own benefit.

An opposite approach to suggesting that empathic people always mobilize themselves toward helping other people in the wake of their empathy is offered by Bloom (2016), who argues that empathy offers no promise of helping others. According to Bloom, empathy may be accompanied by inhibitions, impediments and conflicts featuring high levels of sadness and pain that identify with others' pain and prevent empathic people from acting on their empathy. Bloom thus suggests that it is only anti-empathy that imbues acting persons with the power to assist the many. Quantitatively speaking, therefore, anti-empathy (i.e. indifference) may be seen as possessing a utilitarian moral advantage.

I will not presently discuss all the questions raised by this debate. In any case, it is difficult to deny the fact that empathy entails a certain degree of sadness, and as such may cause the acting person (such as a doctor) to project passivity – all in accordance with the degree to which this sadness affects the acting person, who may thus be unable to help others. If, however, the sadness is not too severe, helping others is still within the realm of possibility.

Favor, as the opposite of passive compassion, is not necessarily driven by sadness and thus – by virtue of being a reasoned act – reflects activity rather than passivity. It should be emphasized that a sadness guided by compassion contains no element of malice, which is why we seek to resolve the factor which evokes compassion within us. This begs the question of how we are to resolve this factor when sadness harms our ability to act – as we have seen above and in our discussion of E.3: 27, c. 2 – 3. Spinoza's remedy is thus meant for transforming emotional chaos into rational order that would drive sadness away and would permit useful activity. Indeed, indifference as a result of self-awareness and reflection is not apathy, but rather a mental tool that places thought before emotion. In other words, it acts as a kind of internal growth engine driving activities that take up the place of the kind of passive sadness that exerts some degree of suppression on our ability to act.

6. *Reasoned activism and implementation*

In E.4: 59 Spinoza proposes the remedy for the emotion of sadness, stating that "In the case of all actions to which we are determined by a passive emotion, we can be determined thereto by reason without that emotion".

As noted above, this remedy is clearly cognitive and is made up of three stages: (1) critical reflection on the emotion (passive); (2) the conscious correction of the error as the key toward (3) corrective reasoned action (active).

In the case of pity, reasoned indifference emphasizes the two elements that must work together: (1) "to endeavor not to be touched by pity", referring to indifference; (2) awareness of the reasons for the special case I have to deal with.

Compassion is non-reasoned on account of being a saddening emotion that suppresses a person's freedom of action. It does not help the giver of compassion or its object. Spinoza takes care to balance our ostensible perception as humans against the need to support and offer assistance in times of needs since it is seemingly inhuman not to feel pity. This is why favor without compassion does not oppose reason so long as it is not carried out as a result of sadness and why it is a reasoned and helpful activity. The person who acts favorably but not as a result of sadness (compassion) is not sad and thus us more helpful than the person who acts as a result of pity. Spinoza's approach in this respect rests on E.3, d. 8 in the *Ethics* which states that "by virtue and power I mean the same thing. Virtue, as far as it is related to man is man's very essence, or nature, insofar as he has power to bring about that which can be understood solely through the laws of his own nature". The nature and substance of favor is thus nothing but conduct driven by reasoned activism, which is the high level suitable for actors who seek to maximize themselves and their abilities as part of nature in general and as part of nature's own reason. This ability is common to all human beings as sentient being due to the common cognitive skills all humans share (Nadler 2007: 218). Gilead (1986) considers overcoming the gap between theory and practice – between knowledge and its practical application alongside the cancellation of the dualism between humans' and nature's reasoning – to be the *Ethics*' highest degree of consciousness (453 – 458). Human beings will thus fail to realize their essence as reasoning beings and will harm their self-realization for as long as this duality between metaphysical (indifferent) knowledge and is practice in everyday life persists. This duality is non-indifference, the high barrier that separates knowledge rom its actual application. The higher consciousness offered by Spinoza internalizes the principle of natural indifference and is particularly available – as reasoned thought - to human self-awareness at stressful and distressful times where humans tend toward passivity (wonder). Indeed, such reasoned thought brings down the wall between theory and practice and transforms theoretical knowledge to what Spinoza refers to as "intelligent conduct".

I believe these abilities can be inculcated by the education system in the form of curricula and self-help training programs. Just as we instruct a child to ignore another child who is bothering him, Spinoza leads us along the paths of *Ethics* to keep away the causes of sadness through reason. In both cases, the correction of ignoring the cause leads us to an adaptation.

7. *On the recognition of good and evil*

The relations between good and evil as well as joy and sadness are not objective determinations that accord with nature, that is indifferent to our fates, but rather subjective determinations that relate to a subject's own inclinations.

In the preface to part 4 of the *Ethics*, and as in E.1: 16, Spinoza emphasizes the fact that nature does not act intentionally or purposefully: nature (or god) acts according to its present imperative, i.e. as a result of its own nature as reality. The role of human reason is to internalize this fact and to equalize our subjective wishes and inclinations with the objective conditions of reality. The subject considered what we desire as good. But this is not the case. Spinoza stresses that we must reverse this order and determines that the good which we desire is desired because it is objectively good and not because we desire it or wish for it. We must thus place thought before emotion once more. Indeed, reason dictates that what reduces our ability to act is bad and what empowers it is good since it accords with our substance as thinking beings. In Spinoza's terms: "An emotion toward a thing which we think of as inevitable [*necessaries*] is more intense, other things being equal, than emotion toward a thing possible, or contingent, that is not inevitable" (E.4: 11), as opposed to "desire arising from pleasure is, other things being equal, stronger than desire arising from pain" (E.4: 18) (my emphases).

The equalization of conditions within a person's awareness represents a balancing of consciousness against reality through reasoned cognitive action – in other words, the use of indifference as a basic state of reason equalizes the conditions of our external reality to our internal awareness according to the necessary chain of causation as it occurred – and this represents as a state of adaptation. What we are concerned with is an informed and reasoned cognitive action which gives rise to a basic situation that ignores images which violate the balance between the real situation and the emotions arising from an unreasoned image of reality. As stated by Spinoza in E.4: 27: "we know nothing to be certainly good or evil except what is really conducive to understanding or what can hinder understanding". In other words, and as Elliott (2017) suggests, "[T]he intellect has knowing good and evil insofar as it has knowledge of the *conatus* [the aspiration of self-preservation] particularly as it pertains to one's either becoming more or less like the rational exemplar [of a rational human being as part of nature]" (266).

Another key sentence in E.4 supports the above: "knowledge of evil is inadequate knowledge" (E.4: 64). This is because recognition of evil represents sadness itself, and as such is beyond a lesser perfection that does not match human substance and nature. It therefore follows that it is an emotion that has nothing

to do with adequate ideas and is an uninformed form of awareness. It also has nothing to do with self-awareness and a reflexive critique of knowledge according to the Principle of Sufficient Reason (PSR) (Schneider 2014). Bennet (1984) distinguishes between two types of reasoned perceptions in Spinoza's thought. The first type, expressed in Part 2 of the *Ethics*, is a guiding idea that settles for the Principle of Sufficient Reason (PSR), while the second type, expressed in Part 4 of the *Ethics*, is an activist internal factor that can withstand external influences on human impulses (184).

Anything that hinders awareness is bad since it misleads us and diverts us to act in a manner that does not accord with our (reasoned) nature. An inadequate awareness lacking the PSR would lead us to perceiving the event or object in question as coincidental rather than necessary – which would then lead to the emergence of sadness in the form of pity, anger, jealousy, revenge and many other causes of sadness and reduce our capacity for action as discussed in E.3: 11 – 13. The lack of an ability to act properly is, as noted above, the result of a lack of balance between the emotive subject and the objective state of nature and constitutes a distorted perception of reality.

In this respect, Spinoza (1976 [1677]) argues that fictions are confused ideals that form an associative connection between distinct images that lack a mutually coherent, logical, determinist and total affinity, a passive kind of view […] that does not involve deliberate reasoning but rather a false connection between the various components of the situation experienced by the acting person.

However, a person free of emotions who lives by the dictates of reason should prove indifferent to this kind of negative emotions, and is not even guided by a fear of death – which is something that she or he does not consider at all. Such a person aspires to the good which accords with her/his human nature – to act, to live, and to maintain her/his being. Spinoza summarizes the preface and the course of discussion on the indifferent contexts of good an evil in this part of the *Ethics* in E.4: 68, p. and s.: free people do not visualize – so long as they are free – any concept of good and evil. A person with adequate ideas flows with nature and neutralizes the concept of evil. She or he will avoid any emotion of sadness including pity, anger, jealousy etc. The connection between humans and the general order of nature is thus the ability to internalize the metaphysical principles suggesting that anything which occurred had been inevitable, since it did in fact occur. In other words, we are concerned with a kind of sublimation: a change of state from inadequate knowledge whose sense impressions lead to sadness to adequate knowledge which ascribes events to necessary causation. The recognition that nothing in nature is intended to make things better or worse or to give rise to anger or

sadness increases the knowing actor's capacity for action. This is because it removes that non-indifferent partition between humans, nature, and nature's methods of action. This reasoned sublimation represents a kind of imitation of indifferent nature and it is the kind of state that should guide our perceptions as subjects. When encountering situations where people are capable of helping themselves and others, there is no cause for isolating human nature and separating it from nature in general which acts by necessity. People should rather flow with nature without erecting any kind of emotional barriers. As stated by Spinoza: "the virtue of a free man is seen to be as great in avoiding dangers as in overcoming them" (E.4: 69).

Spinoza's guidance in this respect may be summarized by the following quote:

> Human power is very limited and is infinitely surpassed by the power of external causes, and so we do not have absolute power to adapt to our purposes things external to us. However, we shall patiently bear [adaptation] what happens to us that is contrary to what is required by consideration of our own advantage, if we are conscious that we have done our duty and that our power was not extensive enough for us to have avoided the said thing, and that we are a part of a whole of Nature whose order we follow (E.3, app. 32. My emphasis).

The things speak for themselves when Spinoza offers us a direct philosophical remedy to our psycho-physical drawbacks centuries before the adaptation and sublimation psychological mechanism offered by Inhelder & Piaget (2019), Vaillant (2000) and others as a remedy to stressful situations, crises and emotions that make it difficult for us to cope with a changing reality.

This mechanism advances the actor's ability to control her or his environment and imbues her or him with a sense of autonomy in decision making and in everyday conduct to the point of self-satisfaction (Ryff & Singer 1998). While the psychological mechanism's notion of sublimation offers some degree of transition from passivity to activity, especially during times of embarrassment or crisis, the internalization of the model proposed in the present article facilitates an educated way of life that is at the actor's disposal at all times, and that is available during times of stress and embarrassment since its application manifests in an educated way of life as a form of continuous conduct rather than a tool to be used for putting out the occasional fire. The psychological mechanism offers the sublimation of tolerance and repression, the blurring of emotions through avoidance, humor and other means that do not arise from the domain of reason. Spinoza's model, on the other hand, offers a reasoned adaptation and the sublimation of a leap from situations of emotive nebulosity to a lucid and clear reasonability according to the perspective of totality (i.e. according the indifferent perspective of Nature as a whole).

8. *Reasoned sublimation in irregular situations*

In the context of education Zourabichvili (2012) uses the term "indifferent" as an expression of the student's transfer to a personal experience, anonymity, and a sense of a general nature such that this feeling will intensify the activity of self-existence. In this spirit, Zourabichvili supports O'Donnell (2018) in that:

> […] this student self will tend to experience herself increasingly depersonalized and decentered as she comes to feel and understand that she is a part of nature, but this, curiously enough, permits of an intensified experience of existence, an openness, and a sense of one's singularity beyond visceral habits and clichés of existing. The movement of decentering and depersonalizing strangely moves us not toward anonymity but toward 'thisness' or *haecceity—this* life (824).

In E.5: 3, Spinoza states that "a passive emotion ceases to be a passive emotion as soon as we form a clear and distinct idea of it", and proceeds to state that "there is no affection of the body of which we cannot form a clear and distinct conception" (E. 5: 4).

Sublimation is a metaphor – a concept borrowed from the material domain and placed in the spiritual domain. The transition from state A to state C is defined as sublimation – a direct transformation from solid to gas, a "sublimation" that bypasses an intermediate stage. Zourabichvili (2002) and Sévérac (2018) use the term "transformation" with respect to Spinoza's educational context. The concept expresses a gradual transition from the improvement of physical qualities and the acquisition of knowledge from childhood to adulthood at the peak of which the adult is released from his self-passive prison, which is influenced by external factors and transformed into action by his own activity.

Our human limitations do not permit us to recognize all the particulars of the chain of causation leading to the event that evoked negative emotions within us. However, as sentient beings who have internalized the metaphysical aspects of nature's determinist causation we do possess the sublimatory ability to perform a reasoned and direct leap from the event as a cause of these negative sublimatory emotions to the cause of causes – the self-cause – nature. Skipping the items of the chain of causation, which forms the intermediate layer between our situation and the cause of causes, becomes irrelevant to our emotional state as soon as we become aware of the initial cause of our situation. This activity is referred to as emotive "refinement", "devolvement", or "the purification of materiality". Reasoned sublimation for the refinement of passive mental states would thus lead us to reasoned adaptation, a basic and profound remedy particularly suitable to us as sentient beings.

9. *The model*

The law of self-preservation (*conatus*) is a general natural law like the physical laws of nature. This law states as follows:

1) Any object whatsoever, and any person in particular aspires to preserve its existence to the best of its ability, and it is only external factors (accidents, wars, bad neighbors, wild animals, viruses, air pollution, radiation, oxidation, other causes of old age and disease and more that may bring its existence to an end.
2) People are finite and delimited beings compared to nature, which is infinite.
3) There always are and always will be other factors with superior abilities that may harm or cancel a person's existence.
4) Human beings care [are not indifferent to] about what occurs within them and in their environment.

Therefore, human beings determine what is good and what is bad according to what they perceive as things that advance or inhibit their existence. Therefore, if I think of something as good and I desire it since it advances my existence or improves my standard of living, I shall act – insofar as it is within my power – to remove anything which inhibits my existence or harms my standard of living.

Since it is human nature to determine what is good and what is bad according to the aforementioned law of self-preservation, human beings classify good and bad as two poles. The mandate of reason dictates that it is better to be located as close as possible to the pole that advances human existence – which is perceived by reason as being good.

It has thus far been possible to summarize and state that the mandate of reason dictates that:

> Anything that empowers my ability to act (the activeness of my being) is good, and anything which restricts my ability to act (the passivity that harms my existence) is bad.

As a thinking and sentient being, humans acting in accordance with the mandate of reason advance themselves toward the good and avoid the bad and thus realize their essence as an existing 'is' in the spirit of the law of self-preservation.

THE GOOD: ACTION	THE BAD: BEING ACTED ON
Self-action	Passivity on account of external factors
Anything that expands or develops my ability to act such as self-fulfillment, happiness, love, sympathy, hope, confidence, cheerfulness, co-operating with others, etc.	Anything that restricts my ability to act such as sadness, depression, jealousy, anger, compassion, shame, apprehension, fear, meekness, etc.

QUESTION: How can I move toward the good side when I am on the bad side (e.g. when I am sad, angry, jealous or depressed)?

ANSWER: By availing yourself of reason, whose role is to internalize the fact that it is only external factors that turn you toward the bad side [the red side]. You can move away from this emotional state (which causes you to be non-indifferent and caring toward the external environment) to an understanding that an indifferent-reasoned state on your part expresses a state of emotional non-submission to the external reasons that suppress the existential activity of your own self. This understanding is kind of an internal activism in and of itself on the part of your own self. When you gaze upon your own self and its ability for action (its activeness), you experience a state of satisfaction that constitutes a kind of self-fulfillment and an increased ability to act on your part which forms the beginning of your movement away from sadness and towards happiness [the blue side].

This movement from the red side to the blue side constitutes an emotional adaptation to the external environment – whatever it may be.

QUESTION: How may feelings of sadness be suppressed, and reasoned thought applied, in order to move from sadness toward feelings of happiness?

ANSWER: This requires the internalization of a number of metaphysical rules that accord with the rationale of all sentient beings (humans). When you are on the red side, these reasoned rules are at your disposal for use as tools for overcoming your sadness and for moving from a passive state toward actual action.

QUESTION: What are these rules?

ANSWER:

1) Nature as an infinite totality is certainly its own cause and the primary and necessary cause of anything that takes place in the universe and necessarily the cause of any thing or object as well as the cause of anything that happens to you on account of external factors.
2) Anything that happens in the world happens necessarily and not ac-

cidentally. There is no coincidence, only our inability to recognize the entire sequence of the chain of causation from proximal cause to event (that may be known) and ultimately to the cause of causes, which is nature as a whole.

3) Human beings are part of nature, and not walled islands within nature. Human beings are fed by nature, breathe in nature, exist in nature, and live within the framework of closer and more distant human groups.
4) The laws of physics – as the laws of nature – apply to human beings just as they apply to anything else (inanimate objects, plants, and living beings). The same applies to the laws of psychology (emotions, thoughts, and sensations) that derive from the law of self-preservation.
5) Nature as a whole is indifferent to human existence and fate as well as to humans' psychological and physical state – whether on the blue or red side – and it conducts itself according to its nature with or without human intervention. The only thing nature cannot control is human reason given that it is a copy of nature's own reasoning.

The role of humans as those who are meant to fulfill the law of self-preservation which is in fact the essence of their existence – mimicking nature and striving to neutralize the effect of external factors on negative emotions, i.e. being emotionally indifferent to anything that inhibits humans' ability to act – given that such inhibition goes against human nature.

QUESTION: How can emotions that suppress the human ability to act and drag them toward sadness and passivity be neutralized?

ANSWER: Once humans understand, know and internalize that the cause of causes, which is god or nature is the primary and necessary cause of human sadness, leading them to skip any attempt to recognize the chain of causation from the saddening event to its primary cause (source) – and thus move from the red side to the blue side (from the heat of emotional fire to the water that douses the emotional flame).

This transition is similar to physical sublimation – the transition from a certain state of matter to another state of matter while skipping all intermediate states. Humans adapt to their environment and flow with the river (the world) and do not attempt to swim against its strong current. Humans adapt to their environment and thus begin to act in a reasoned matter, availing themselves of the current (the conduct of nature) and directing themselves to reasoned conduct – meant for advancing their self-preservation – their existence as part of the world rather than a separate (non-indifferent) part of it.

10. *Epilogue*

Reasoned sublimation as a "clear and distinct idea" transitions us from "passive emotion" to a "clear and distinct conception of the situation" without the mediation of an intermediate layer in the form of the chain of causation that led to this situation.

In the context of the present discussion, the sensation of pity ceases to be passive when we frame it as a lucid, clear, and adequate idea, which is what it actually is from the perspective of nature (or god) as a whole. Viewing the adequate idea is a reasoned activity of "condition equalization" between the subjective emotion of sadness as a confused idea and the objective natural state which is indifferent to the object of our sadness. "Condition equalization" thus refers to the "framing" of the emotion in a reasoned-metaphysical framework, its neutralization with respect to the subjective environment, and its perception within the framework of the objective (true) state of nature. In other words, it refers to an understanding that we are concerned with a situation caused by necessity (Delassus 2014). The transition from the imbalanced state of reduced perfection to a balanced state of extensive perfection is a reasoned state of copying or a reasoned imitation of nature's indifferent conduct. In other words, what we are concerned with is the self-awareness of a person's place and objective state within an infinite totality. Indeed, Spinoza himself provides an imprimatur for this understanding in E.5: 6: "Insofar as the mind understands all things as governed by necessity, to that extent it has greater power over emotions, i.e. it is less passive in respect of them".

This is because nature itself is free of emotions, does not conduct itself according to human concepts of good and evil, does not sense the emotions of either joy or sadness, and does neither love nor hate any human whatsoever (E.5: 17 & 17 c.). And again, in order to move from passive to active, Spinoza guides us, as we instruct a child to ignore a child who is bothering him. An adaptive child is an adult child, an adaptive person is an intelligent person.

Abraham Mounitz
drmounitz@gmail.com
Zefat Academic College

References

Barbaras, F., 2007, *Spinoza: La Science Mathematique du Salut [The Mathematical Science of Salvation]*, CNRS Editions, Paris.

Bennet, J., 1984, *A Study of Spinoza's Ethics*, Cambridge University Press, Cambridge.

Bennet, J., 1996, "Spinoza's Metaphysics", in D. Garret (ed.), *The Cambridge Companion to Spinoza*, Cambridge University Press, Cambridge: 61-88.

Ben-Shlomo, J., 2012, *Ha'Etgar Shel Ha'Spinozism [The Challenge of Spinoza and Spinozism]*, Carmel Publishing, Jerusalem [Hebrew].

Bloom, P., 2016, *Against Empathy: The Case for Rational Compassion*, HarperCollins Publishers, New York, NY.

Curley, E., 2014, "How Spinozistic was Toland's Pantheism", in A. Kasher (ed.), *Models of God and Alternative Ultimate Realities*, Springer Science and Business Media, Dordrecht: 643-653.

Della Rocca, M., 1996, *Representation and the Mind-Body Problem in Spinoza*, Oxford University Press, Oxford.

Della Rocca, M., 2008, *Spinoza*, edited by B. Leiter, Routledge, London-New York.

Delassus, E., 2014, *Être malade: accepter pour résister* [Being sick: accept to resist], visited 01 June 2019, https://hal.archives-ouvertes.fr/hal-01082013/document.

Delassus, E., 2018, *Spinoza - La troisième partie de* L'Éthique*: Une géométrie des affects* [Spinoza - Part III of the *Ethics*: The Geometry of Emotions], visited 01 June 2019, https://hal.archives-ouvertes.fr/hal-01710193/document.

Elliot, J., 2017, "Maimonides and Spinoza on God and Evil: A Reappraisal of W.Z. Harvey", in *Iyyun: The Jerusalem Philosophical Quarterly*, 66, 3: 258-270.

Gilead, A., 1986, *Darka Shel Torat Shpinoza Le'Shita Filosofit [The Way of Spinoza's Philosophy toward a Philosophical System]*, Bialik Institute, Jerusalem [Hebrew].

Glare, P.G.W., 2012, ed., *Oxford Latin Dictionary*, Second Edition, Oxford University Press, Oxford.

Golomb, J., 2015, "Kama Refleksiyot Bi'Sdot Ha'Krav Shel Ha'Filosfiya: Shirat Ha'Yona Le'Regel Tsetiy Le'Gimlaot [Reflections on the Battlefields of Philosophy], in *Iyyun: The Jerusalem Philosophical Quarterly*, 64, 4: 451-465.

Green, K., 2016, "Spinoza on Self-Hatred", in *Iyyun: The Jerusalem Philosophical Quarterly*, 65, 1: 73-95.

Inhelder, B., Piaget, J., 2002, *The Growth of Logical thinking From Childhood to Adolescence*, Routledge, London and New York.

Kisner, M.J., 2008, "Spinoza's Virtuous Passions", in *The Review of Metaphysics*, 61, 4: 759-783.

LaBuffe, M., 2010, *From Bondage to Freedom: Spinoza on Human Excellence*, Oxford University Press, Oxford.

Mounitz A., Berenson. A., 2016, "The Human Desire for Pleasure as a Basis For the Unity of Mankind: Empiric Research Validating Spinoza's 350 Year-Old Philoso-

phy", in *European Journal of Education Studies*, 2, 3: 64-79.

Nadler, S., 2011, *A Book Forged in Hell: Spinoza's Scandalous Treatise and the Birth of the Secular Age,* Princeton University Press, Princeton, NJ.

O'Donnell, A., 2018, "Spinoza, experimentation and education: How things teach us", in *Educational Philosophy and Theory*, 50, 9: 819-829. doi:10.1080/00131857.2017.1382351.

Piaget, J., Inhelder, B., 2019 [1966], *La Psychologie de l'Enfant*, Presses Universitaires de France, Paris, Eng. Tr. by H. Weaver, 2019, Basic Books, New York, NY.

Rutherford, D., 2010, "Spinoza's Conception of Law: Metaphysics and Ethics", in Y. Melamed and M. Rosenthal (eds.), *Theological-Political Treatise*, Cambridge University Press, Cambridge: 143-167.

Ryff, C.D., Singer B., 1998, "The Contours of Positive Human Health", in *Psychological Inquiry*, 9, 1: 1-28.

Schneider, D., 2014, "Spinoza's PSR as a Principle of Clear and Distinct Representation", in *Pacific Philosophical Quarterly*, 95, 1: 109-128.

Sévérac, P., 2018, *L'enfant est-il un adulte en plus petit? Anthropologie et psychologie de l'enfance à partir de Spinoza* [Is the child a smaller adult? The Anthropology and Psychology of Childhood according to Spinoza], visited 01 June 2019, https://journals.openedition.org/asterion/3406

Spinoza, B., 1976 [1677], *Tractatus de Intellectus Emendatione*; n.p., Amsterdam. Heb. tr. by N. Spiegel, introduction and notes by J. Ben-Shlomo, 1976, *Ma'amar 'Al Tikun Ha'Sechel*, Magnes Press, Jerusalem.

Spinoza, B., 2002[1677], *Tractatus de Intellectus Emendatione*; n.p., Amsterdam. Eng. Tr. by S. Shirley in *Spinoza: Complete Works*: 1-30, 2002, *Treatise on the Emendation of the Intellect*, Hackett Publishing Company, Inc., Indianapolis, IN.

Spinoza, B., 2002a[c. 1660], *Korte Verhandeling van God, de mensch en deszelvs welstand*; n.p., Amsterdam. Eng. Tr. by S. Shirley in *Spinoza: Complete Works*: 31-107, 2002, *Short Treatise on God, Man, and His Well-Being*, Hackett Publishing Company, Inc., Indianapolis, IN.

Spinoza, B., 2002b[1677], *Ethica Ordine Geometrico Demonstrata*; n.p., Amsterdam. Eng. Tr. by S. Shirley in *Spinoza: Complete Works:* 213-382, 2002, *Ethics*, Hackett Publishing Company, Inc., Indianapolis, IN.

Spinoza, B., 2002c[1661-1676], *Epistolae*; n.p., Amsterdam, Rijnsburg, Voorburg, The Hague. Eng, Tr. by S. Shirley in *Spinoza: Complete Works*: 755-960, 2002, *The Letters*, Hackett Publishing Company, Inc., Indianapolis, IN.

Steenbakkers, P., 2009, "The Geometrical Order in the Ethics", in O. Koistinen (ed.), *The Cambridge Companion to Spinoza's Ethics*, Cambridge University Press, Cambridge: 42-55.

Vaillant, G.E., 1993, *The Wisdom of the Ego*, Harvard University Press, Cambridge, MA.

Vaillant, G.E., 2000, "Adaptive Mental Mechanisms: Their Role in a Positive Psychology", in *American Psychologist*, 55, 1: 89-98.

Zahavi, D., 2015, "Empathy ≠ sharing: Perspectives from Phenomenology and Developmental Psychology", in *Consciousness and Cognition*, 36: 543-553.

Zourabichvili, F., 2002, *Le conservatisme paradoxal de Spinoza: Enfance et royauté* [Spinoza's Paradoxical Conservatism: Childhood and Royalty], Presses Universitaires de France, Paris.

Zourabichvili, F., 2012 [1994, 2003], *Deleuze: Une philosophie de l'événement*; *Le vocabulaire de Deleuze*, Presses Universitaires de France, Paris; Ellipses, Paris; Eng. Tr. by K. Aarons, Edited by Gregg Lambert and Daniel W. Smith, *Deleuze: A Philosophy of the Event together with The Vocabulary of Deleuze*, Edinburgh University Press, Edinburgh.

Appendix 1
Sample Questionnaire Based on Spinoza's Model

INDEPENDENT VARIABLE: Intellectual Capacity.

HIDDEN INDEPENDENT VARIABLE: The extent of applied control of intellectual knowledge

DEPENDENT VARIABLE: Degree of Adaptability to Stressful Situations (application capacity – the reasoned transition from non-indifferent emotion to indifferent reason).

MEASUREMENT SCALE: (1) – to a very great extent; (2) to a great extent; (3) to a medium extent; (4) to a very limited extent; (5) not at all.

QUESTIONS:

1. To which extent do you believe that whatever occurs in the world occurs necessarily rather than coincidentally?
2. To which extent do you believe that what has taken place in the world thus far is the result of circumstantial conduct that took place with or without you?
3. Do human beings, in your opinion, form part of nature as a whole (the world) or are they a separate 'nature reserve' (an island) within nature as a whole?
4. In your opinion, is nature as a whole (the entire world) indifferent to our fate as human beings?

   No need to continue the questionnaire if question 4 is answered in the negative.
5. To which extent can the recognition of nature's indifference to your predicament in stressful situations make it easier for you to adapt to situations you find unpleasant (stressful situations)?
6. To which extent does this understanding (presented above in questions 1-4) allow you to neutralize negative emotions such as sadness, anger, jealousy, pity, and embarrassment?
7. To which extent do you view the neutralization of the aforementioned emotions as something that might help you act in stressful situations?

8. To which extent do you view the neutralization of negative emotions (detailed in question 7) as an empowerment of your personal capacity to act better in stressful situations?
9. To which extent, in your opinion, shall the aforementioned recognition advance your capacities to help others (such as a traffic accident casualty)?
10. To which extent do you believe that an internalization of the understanding that stressful situations are necessarily caused (i.e. are the result of a chain of causation that takes place with or without you) is a practical activity rather than a form of theoretical thinking?

Appendix 2
Interviews
(The names are fictitious)

SILVIO is 53 years old, married and has one daughter who is currently completing her mandatory military service and one son, currently 12 years old. Lives in the city of Akko (Acre). His cognitive situation is excellent, his memory is excellent, and he is in good physical health. He is a computer programmer by training. He experienced a traumatic event three years ago when he discovered his wife sleeping with another man in his [Silvio's] own bed. His wife subsequently had a restraining order issued against him. His mother became ill a year ago and he dedicated most of his time to caring for his mother at the hospital. His mother literally passed away in his arms. His long absences from work led to his termination. He has not been able to work since the event. When I met him, he was completely destitute and alternated between sleeping at friends' houses or at his married sister's house and eating at their expense. His mental situation was poor; he was incapable of taking responsibility for himself, or even carrying out basic tasks that could improve his living circumstances. Searching for a job, for example, seemed to him like an almost impossible task akin to scaling an impossibly high wall. His mental state appeared to indicate a kind of emotional nullity, total desperation and an inability to employ his good cognitive abilities.

I met him once a week for 3-hour sessions for a period of about three weeks as well as conducted twice-weekly phone calls lasting about 30 minutes each. I gradually exposed him to Spinoza's model in the course of our conversations and stressed the indifferent view which he must apply to the reality surrounding him as well as the external causes that neutralize his ability to act and lead him to surrender to his sadness, depression and tendency to view simple tasks as being impossible. Employing the lens of the Spinozan model I presented to him allowed him to successfully evoke cognitive abilities and reflect on his situation. According to him, "I internalized the principle [of indifference], that what happens around me forms part of nature as a whole that is indifferent to my fate and this [understanding]

imbued me with the ability to view myself [reflexively] in relation to the reality in which I exist and the emotional swamp I am trudging through. As we proceeded, I understood that it was only my own power [abilities] that could sustain me and that my reasoned mind should cope with this morass which is essentially a kind of virtual reality". It did not take long for him to find a new job and he is currently functioning well. His adaptation is not only manifested in his words but also in his behavior – he does not feel anger toward or seeks vengeance on his wife, and also accepts his mother's death as a determinist *fait accompli*. According to him "the wall that formerly existed between me and the task of finding a new job suddenly became a thin thread I could easily skip over".

Dina is a 62-year-old married woman with two daughters and 5 grandchildren. Dina has been working as a senior nurse at a Northern Israeli hospital for the past 24 years. At the time of our first interview she served as the head nurse of the internal medicine ward, a position she was assigned to after working in other wards including the emergency medicine ward (the ER). Her physical state is very good, she exercises twice a week and her cognitive state is excellent. She has recently been finding it very difficult to care for a severely ill patient which has been admitted to her ward. She states that she "cannot understand how someone like myself, who has acted properly in difficult cases and cared for injured ER patients, is incapable of caring for her and helping her at present". She proceeded to explain that "this patient has been hospitalized in my ward for a long time, she has a charming personality and a big heart, she doesn't complain and does not groan; she captured my heart and I formed such an emotional connection to her that I identify with her difficult situation and it kills me that I can't care for her myself".

It is quite readily apparent that the high functional barrier that arose between Dina and the professional approach required of a longtime and experienced nurse as well as her emotional and non-indifferent identification with the patient did her a disservice and caused her mental distress to the point of losing the ability to realize her professional capacity. This non-indifference caused by her emotional sympathy led Dina to what Spinoza described as pity, which he brands as a kind of sadness that affects a person's capacity for action.

It was as early as our first meeting that Dina revealed herself as possessing an excellent capacity for reflection and self-criticism that accords with reality and yet also as a person who lacks "emotion-neutralizing" cognitive tools for coping with her situation. Five 90-minute instructional meetings allowed Dina to internalize Spinoza's determinist model such that she could use it in practice. She noted "the reasoning abilities I acquired from the model, and especially the transition from an emotional state to the understanding that anything that has occurred happened necessarily [determinism]. Nature as a whole as well as external causes (the patient's severe prognosis) beyond my control caused me to become passive and unable to realize my professional knowledge in an objective manner and in the patient's favor". It was not long before Dina adapted to her objective situation and began caring for the

patient in a more emotionally robust manner that she acquired by placing reason before emotion and by perceiving the bigger (determinist) picture of reality.

ASHER is a 36-year-old man, married and without children. He is good-looking, solidly built, well-dressed and brimming with calm self-confidence. My first meeting with Asher, who owns a successful Tel Aviv restaurant, took place on the steps of the Tel Aviv Rabbinical Court where he was waiting to be called for another deliberation of the divorce he filed against his wife. I accompanied him on the steps and to the Courtroom door.

Our background discussion of his case revealed an emotional person full of anger and feelings of vengeance. "Let me just finish these divorce proceedings and I'll make sure that she smells the flowers from six feet under", he said defiantly in the direction of his wife who was also waiting for the deliberation nearby. This ostensibly appears to be a complex case of a wounded ego that almost got into a fight with strangers whose own deliberations were scheduled before him.

The deliberation with the rabbinical judges ended with no practical solutions. I had since accompanied him alternately for three and a half months including a one-month break. The meeting protocols reveal 13 one to two-hour sessions whose frequency and length was determined by his priorities and by business affairs that required his personal attendance at his restaurant. I spent time observing his walking during the restaurant's opening hours and conveyed my critical remarks hoping to make him view himself from the outside. It was important to me that he accustomed himself to reflective observation such that he may be able to crack the cover of blindness his ego developed between himself and reality. The first signs of such a crack appeared after he initiated a one-month break in our sessions. Our sessions resumed after this one-month break. At this point in time, he was much softer and more attentive.

The 'philosophical therapy' in the spirit of the Spinozan model took place very intensively and was accompanied by guided and attentive reading of key passages in Spinoza's *Ethics* on his part (as a distrustful person, he had to examine the source to the therapeutic method. This also built up a degree of trust between me and him and convinced him that I was not experimenting on him).

The results exceeded our expectations. Asher stated that "I was reborn. I am a different person. Our reading and your explanation allowed me to view reality with rose-tinted lenses and neutralized a lot of emotions in me. In hindsight, I realize that I was imprisoned by my ego and acted as if the devil was pulling my strings and manipulating my actions. I now know where true power lies. I used the reasoned indifference I internalized to cut the devil's strings and I view my actions as driven solely by my internal generator and not by any external factor".

In summary, Asher is no longer angry toward his wife, is no longer jealous, and is completely without feelings of vengeance. He states that he accepts her as she is – as a necessary thing whose character, desires and behavior are not his responsibility.

In Asher's words, "there are reasons for this that not even she recognizes, and I don't blame her for that". His impressive sublimation ability that allowed him to transform from a calm state to an angry, threatening and violent state in a heartbeat was transformed through his internalization of Spinoza's philosophical model into a reasoned kind of sublimation, an adaptation to reality. The same reality that was formerly stormy and tempestuous appears – when viewed reasonably – as something that is not coincidental but rather necessary, as being derived by a chain of causation, and as something that must be accepted calmly since human actions form part of the necessary conduct of nature as a whole. Human beings are thus not a separate realm of nature but rather part of nature as a whole acting according to natural laws. Therefore, anything that occurs was derived from a chain of causation that began with the six days of creation.

Integrating first and second nature: Rethinking John McDowell's liberal naturalism[1]

Thodoris Dimitrakos

Abstract: McDowell's 'naturalism of second nature' is one of the most important attempts to defend liberal naturalism in contemporary Anglo-American philosophy. Liberal naturalism stands as an umbrella term for philosophical accounts which place normative entities within the realm of nature. McDowell's attempt to form an account which leaves room for normativity in nature is based on the distinction between 'first' and 'second' nature. In the present paper I shall attempt to shed light on McDowell's notions of 'first' and 'second' nature and thereby provide an account about the status of normativity which is McDowellian in spirit. However, I suggest that human sciences offer an even more acute challenge to the conceptions of nature that aspire to be liberal naturalist, and I argue that McDowell's account needs to be completed with a further conceptual distinction in order to cope with the challenge of human sciences. In particular, I argue that we should distinguish between the notions of explanatory reduction and normative eliminability.

Keywords: liberal naturalism; normativism; John McDowell; Second Nature; Hegel.

1. *Introduction*

In contemporary Anglo-American philosophy, "liberal naturalism" stands as an umbrella term for the philosophical accounts which seek for logical room and, respectively, for a viable intermediate position between scientific naturalism and supernaturalism (Macarthur & DeCaro 2010: 9). Scientific naturalism qua 'naturalism' identifies reality with the realm of nature and qua 'scientific' identifies the realm of nature with the subject matter of the natural sciences. Supernaturalism, on the other hand, is taken to be the philosophical commitment to the existence of "entities or qualities or relations of a very strange sort, utterly different from anything else in the universe" (Mackie 1977: 38). Philosophers who are committed to those entities or relations are also prone to accept a special kind of epistemic faculty (Macarthur & DeCaro 2010: 3) – such as mental intuition – which is supposedly indispensable for grasping supernatural entities

[1] I am indebted to Evgenia Mylonaki for our long-term discussions on naturalism.

philinq VIII, 1-2020, pp. 37-68
ISSN (print) 2281-8618-ETS DOI: 10.4454/philinq.v8i1.216
Submitted: June 2018
Accepted: June 2019

or relations. Nowadays very few philosophers would accept for themselves the label of supernaturalist. Most philosophers place themselves within the broader philosophical camp of naturalism, thus, the term supernaturalism is used more as a philosophical accusation rather than a label which someone would endorse.

As many surveys of contemporary Anglo-American philosophy imply (Kitcher 1992: 54; Leiter 2004: 2; Rorty 2010: 57), most philosophers who reject supernaturalism adopt a scientific version of naturalism. These philosophers think that metaphysical naturalism (Risjord 2014: 9), i.e. the view that the human mind is part of the natural world, necessarily entails epistemological or methodological naturalism, i.e. the view that natural scientific understanding is the only legitimate way of making things in the natural world intelligible. Furthermore, many of them think that metaphysical naturalism also entails meta-philosophical naturalism (Risjord 2016: 2), i.e. the view that philosophy is continuous with the empirical science. Despite its popularity, scientific naturalism does not come without philosophical worries. The main problem is related with what is often called the "normative problem" and concerns the status of normativity. "Normativity concerns what we should or ought to do and our evaluations of things or states of affairs" (Macarthur & DeCaro 2010: 1). Normative facts cover the entire range of human consciousness and behaviour, presuppose the notion of rationality, and are characterized by dichotomies such as true/false (epistemic facts), good/bad (practical facts), beautiful/ugly (aesthetic facts), and so on. According to scientific naturalism, normative facts should be explained in the way that all other facts are explained, i.e. by employing the explanatory patterns of science.[2] This entails that normative explanations are not genuine explanations and should be reduced to descriptive or empirical explanations of science. In short, we could say that scientific naturalists are anti-normativists (Turner 2010).

Rejecting the genuineness of normative concepts threatens to turn notions like freedom, responsibility, morality, justice, or even truth into mere illusions. Against this threat many contemporary philosophers seek for an alternative which still rejects supernaturalism but leaves room for normative facts in the realm of nature. This alternative is often called "liberal naturalism" in contrast to the restrictive conception of scientific naturalism. The main position of liberal naturalism is that embracing metaphysical or ontological naturalism does not necessarily entail epistemological or meta-philosophical naturalism. Thus, normative concepts can unproblematically be used in order to make things in

[2] In Stephen Turner's terms (2010: 11), so-called normative facts should be placed into the "ordinary stream of [empirical] explanation". This means that they should be reconstructed out of conceptual materials that belong to science. See also §2 below.

the natural world intelligible. In this sense liberal naturalists are normativists.

One of the leading figures among contemporary liberal naturalists is John McDowell (Leiter 2004; Macarthur & DeCaro 2004; 2010; Rorty 2010; Turner 2010). McDowell uses a variety of terms in order to express the idea that there is room for normativity in the ontological territory of nature: "naturalism of second nature" (MW[3]: 86), "relaxed naturalism" (MW: 89), "naturalized Platonism" (MW: 91), "Aristotelian naturalism" (McDowell 1998a: 197), and of course "liberal naturalism" (McDowell 2009: 262). McDowell's strategy for coping with the normative problem is explicitly influenced by Wittgensteinian quietism[4] and does not aim at the construction of a theory that can resolve the problem, but at the detection of the ideas or assumptions that create the 'anxiety' in the first place. In the process of dissolving the problem of placing the normative into the realm of nature, McDowell exploits the Aristotelian notion of second nature. Thus, he ends up by rejecting the scientistic assumption that nature is exhausted by the subject matter of the natural sciences and by accepting the genuineness of the normative realm.

In the present paper, I shall attempt to forge an account which sheds light on McDowell's notions of 'first' and 'second' nature and thereby provide an account about the status of normativity. This account aspires to be liberal naturalistic, and furthermore McDowellian, with respect to three fundamental issues: a) it does not depict normative facts as illusions and does not consider normative vocabulary eliminable; b) it retains the distinction between the kind of intelligibility which is proper to reason and other kinds of intelligibility and thus includes normative explanations in a proper understanding of the world; and c) it understands normativity in a collectivistic and historical way in contrast to ahistorical and individualistic philosophical accounts. On the other hand, the account I am attempting to forge is distanced from McDowell's line of thought since mine aims to reply to the question about the relation between first and second nature in a more constructive – in contrast to a quietist – way. In addition, my account ends up being more sympathetic to, and reconciled with, the empirical sciences in general and the human[5] (social and behavioral) sciences in particular.

[3] The abbreviation MW stands for *Mind and World* (McDowell 1996).

[4] According to Leiter (2004: 2-3), most philosophers who react to Quinean-inspired naturalism are influenced by Wittgenstein. Thus, the two main camps with regard to the metaphysical worries over the relation between reason and nature are the naturalism of Quinean descendance and Wittgensteinian-inspired quietism. As also Richard Rorty (2010: 57) points out this division reflects "the deepest and most intractable difference of opinion within contemporary Anglophone philosophy".

[5] I am not following here the standard but quite obscure academic division between the social sciences and humanities. With the term 'human sciences' I refer to all empirical sciences that are concerned with human thinking and behaviour such as sociology, economics, social anthropology, cognitive science, etc.

More specifically, my line of argument will take the following course: in the following section I show how we can understand the normative problem as an explanatory problem or a problem of kinds of intelligibility. In the third section I present McDowell's liberal naturalist account as it was presented in MW. In the fourth section I present a popular anti-normativist argument which I call the 'transition problem' and argue how we should deal with it with the aid of the McDowellian work. In the fifth section I refer to McDowell's revisions of his initial perspective. Next, I argue that while the revisions made McDowell's account even more refined and science-informed, the latter does not respond with the same sensitivity to the lessons from the human sciences. Thus, I suggest that the human sciences offer an even more acute challenge to the conceptions of nature that aspire to be liberal naturalist. In the sixth section I present three available philosophical strategies to avoid scientific naturalism in light of the emergence of the human sciences. I show that the first two are problematic and I conclude that only McDowell's strategy is viable. In the seventh section I argue that despite its advantages, McDowell's quietist strategy is not strong enough against the threat of scientific naturalism which takes into account the emergence of human sciences. Thus, I provide a further argument in order to restrain the threat of scientific naturalism and sketch a liberal naturalist account. Finally, I discuss two main consequences of my liberal naturalist conception which diverges gravely from McDowell's view.

2. *The normative problem as explanatory problem*

In recent years the problem of the status of normativity has attracted numerous contributions in various areas of philosophical inquiry (Finlay 2010: 331-333). The source of epistemic norms (Kornblith 1993, Chan 2013), the normative dimension of meaning (Gibbard 2012), the motivational or binding force of practical reason (Parfit 2006, Broome 2007), the special status of social sciences (Risjord 1998; 2014) are only a few examples of coping with distinct philosophical problems which are related to the general problem of the status of normativity. I will focus on the normative problem from a very specific standpoint, i.e. I will treat the normative problem as a problem of explanatory genuineness. From this standpoint the crucial question is the following: should we take normative explanations to be a genuine kind of explanation or we should understand them as merely provisional forms of intelligibility which need to be reduced to the legitimate explanatory patterns of the empirical sciences?

2.1. Normative explanations and empirical-scientific explanations

Normative explanations make things intelligible by showing how they conform to norms. For instance, if someone asks why I believe q, I could reply that I believe p and I also believe that if p then q. In other words, I can explain my belief modification by subsuming it in the epistemic norm of *modus ponens*. Accordingly, if someone asks me why I help a stranger to find her way home, I can reply that helping her is an expression of solidarity and solidarity is something good. In other words, I can explain my action by subsuming it to a practical norm. Empirical-scientific explanations, on the other hand, make things intelligible in a quite different way: they show how they fit in the causal order described by the empirical sciences. These are the cases where I explain someone's belief or action by referring to various psychological or social factors that cause the belief or the action in question. Note that when I explain things in the one way or the other it is indifferent whether I refer to myself or to another person. I can use both kinds of explanation for either myself or for another person.[6] What matters is that normative and empirical-scientific explanations have a distinct logical form. Empirical-scientific explanations make things intelligible by placing them into the causal order (laws of nature, mechanisms, etc.) implied by each scientific discipline. In this sense, explanations of this sort are value-free. Normative explanations, on the contrary, make things intelligible by showing how they conform to values. Explaining a belief or an action in a genuine (as opposed to merely a seemingly) normative way is equivalent to justifying this action or belief. Thus, in this sense, genuine normative explanations are evaluative and can be equated to justifications.

2.2. Normative explanations and justifications

Against this, one could object, as Derek Parfit (2006: 43) does, that "[w]hen we claim that someone's state would be sufficient to explain his doing something, we do not seem to be claiming that this person accepts a justification for doing this thing". Indeed, this distinction should be taken into account in various philosophical debates, like the one about the motivational force of reason which Parfit discusses. However, from the standpoint I want to discuss the problem of normativity here, there is no need for stressing the distinction between normative explanations and justifications. Normative explanations,

[6] Akeel Bilgrami (2010: 25) suggests that "it may be sensible to replace the terms the 'first-person' point of view and the 'third-person' point of view with 'the agent's' or 'the engaged' point of view and 'the observer's' or 'the detached' point of view, respectively)". I think that he has a point because what is crucial here is the difference in the logical form of explanation and not whether I am explaining myself or another person.

like justifications, are acts of reasoning which show how a belief or an action conforms to a norm.

2.3. The normative dilemma

From this standpoint we can express the normative problem in terms of the following dilemma: either normative explanations are not genuine and thus they can and should altogether be reduced to the domain of empirical-scientific explanations (scientific naturalists' programmatic claim), or normative explanations are *sui generis* and therefore cannot be reduced to other kinds of explanations without losing valuable informational content about the world (normativists' claim).[7] Thus, in what follows, I will examine McDowell's distinction between first and second nature primarily as an attempt to rescue the genuineness of normative explanations. I will also attempt to provide an argument in favor of the ineliminability of normative concepts. Treating the normative problem as a problem of explanatory genuineness does not, of course, solve all the problems that are related to the status of normativity (the motivational force of reason, the source of epistemic norms, etc.). However, the ineliminability of normative explanations is the last line of defense of normativism, for no version of normativism is viable without presupposing that at least some normative explanations are genuine. In this sense, it is of crucial importance for every other aspect of the problem of normativity.

3. *Second Nature as a Reply to the Normative Problem*

McDowell in his seminal *MW* faces the normative problem as a problem concerning our intellectual freedom. The question is how can we understand the relation between sense experience and thought. The puzzlement arises because something merely natural, i.e., our sense organs and their function, and something which is traditionally conceived in a normative way, i.e., our concepts, should be conceived in a unified explanatory context. In other words, how it is that merely natural movements like the affection of our sense organs by sensory stimuli can affect movements within the space of reasons like the grounding of a belief.

[7] The central normativist thesis does not necessarily entails that normative facts correspond to an (ontologically) distinct substance. See for instance Donald Davidson's (1980) "anomalous monism", according to which, mental phenomena do possess causal powers but they are irreducible to physical events. Actually, the whole point of liberal naturalism is to preserve the normativist claim without appealing to any kind of ontological dualism.

3.1. Bald naturalism vs rampant platonism

In McDowell's view, this puzzlement is the result of a mistaken idea – widespread in modern and contemporary philosophy – which equates the realm of nature with the realm of natural law.[8] In other words, this idea identifies the ontological territory of nature exclusively with the subject matter of empirical-scientific knowledge and leads to an entrapment between two unsatisfactory choices. The first is "bald naturalism" that rejects the idea that the logical space of reasons is *sui generis*. Bald naturalism is just another label for what we have described as anti-normativism. The programmatic aspiration of this philosophical perspective is to eliminate the normative vocabulary by reconstructing the logical space of reasons "in terms that belong in the logical space of natural-scientific understanding" (MW: xxii). The elimination of the normative vocabulary is the result of the conviction that normative explanations are not genuine explanations and have to be reduced to empirical-scientific explanations. The other choice is "rampant platonism" (MW: 78), which preserves the position that the space of reasons is *sui generis* but at the cost of regarding it as something extra-natural. In short, rampant platonism is a form of supernaturalism. The conclusion is that as long as we equate nature with the object of empirical-scientific understanding, we are forced to decide what idea we should sacrifice: the genuineness (authenticity) of normative explanations, and hence, of the space of reasons, or its position within the realm of nature. However, rather than sacrifice one of those two ideas we should reject the equation between the realm of nature and the subject matter of empirical-scientific knowledge. This would enable us to suggest that human rationality is both natural and *sui generis*;[9] it belongs to the realm of nature but cannot be made intelligible with the use of the conceptual tools of the empirical sciences (MW: 88).

[8] In *MW* (1996: 71, fn 2), McDowell follows Russell and identifies the explanatory patterns of science with the formulation of natural laws. The reason McDowell rejects Rorty's term is that he believes that reasons can be causes. Later he changed his mind about this Russellian conception. I will come back to this in the fifth section.

[9] The programmatic goal of this argument is proclaimed in a single footnote at the beginning of *MW*. There McDowell repeats the famous Sellarsian quote: "In characterizing an episode or a state as that of knowing, we are not giving an empirical description of that episode or state; we are placing it in the logical space of reasons, of justifying and being able to justify what one says" (Sellars 1997: 76); he also adds: "In much of the rest of these lectures, I shall be concerned to cast doubt on Sellars's idea that placing something in the logical space of reasons is, as such, to be contrasted with giving an empirical description of it. But the theme of placing things in the space of reasons is of central importance for me" (5).

3.2. The Aristotelian notion of second nature

McDowell's way of expressing the idea that nature is not identified with the subject matter of scientific knowledge is to suggest that "nature includes second nature" (MW, xx). In other words, he exploits the Aristotelian notion of 'second nature'[10] in order to argue that people are born as mere animals and during the course of their maturation acquire the ability to respond to reasons. According to McDowell's Aristotle, what distinguishes human beings from mere animals is the rationality they acquire through their upbringing, through their *Bildung*.[11] Habit (the Greek ἕξις) and social training in general make individuals capable of acquiring conceptual powers and "[w]hen we acquire conceptual powers, our lives come to embrace not just coping with problems and exploiting opportunities, constituted as such by immediate biological imperatives, but exercising spontaneity, deciding what to think and do" (MW: 115). Therefore, "[o]ur nature is largely second nature, and our second nature is the way it is not just because of the potentialities we were born with, but also because of our upbringing, our *Bildung*" (MW: 84). The process of *Bildung* is characterized by a central element which is the acquisition of a natural language (MW: 125). Human beings are born as mere biological creatures and turn into thinkers and agents, that is, into rational animals, through their initiation of a language. The language does not serve only as a means for communication but also "as a repository of tradition, a store of historically accumulated wisdom about what is a reason for what" (MW: 126). In this sense, rationality has an unambiguous collectivist and historical dimension. Rationality is not a feature of the isolated individual, say the transcendental subject. Rationality distinguishes human beings from mere animals and stems from the acquisition of a language which entails the acquisition of the historical wisdom of a community. This means that rational animals can be this kind of animal only as part of a tradition, viz. as part of a historically shaped community.

To sum up, the notion of second nature can help us create the proper logical space between scientism (bald naturalism), which makes normativity look like an illusion, and ontological dualism (rampant platonism), which makes our capacity to respond to reasons – our freedom[12] – "look like an occult power"

[10] McDowell (1996: 84) admits that the notion is all but explicit in Aristotle's works but he thinks that it can be reconstructed out of the Aristotelian thinking. For a brief history of the notion, see Gubeljic. et al. (2000).

[11] If we generalize the way Aristotle conceives the moulding of ethical character, we arrive at the notion of having one's eyes opened to reasons at large by acquiring a second nature. I cannot think of a good short English expression for this, but it is what figures in German philosophy as *Bildung* (MW: 84).

[12] McDowell follows Kant and generally the tradition of German idealism in identifying freedom with rationality: "When Kant describes the understanding as a faculty of spontaneity, that reflects his

(MW: 83). The rejection of the assumption that the realm of nature is co-extensive with the subject matter of empirical sciences and the division of nature into first and second leads to a liberal naturalist conception that places normative phenomena into the wider territory of nature. According to this conception, the normative explanations are not reducible to the empirical-scientific explanations, for the former are concerned with phenomena within the second nature while the latter with phenomena of the first nature.

4. *Second nature and the transition problem*

One obvious objection to this conception is that the emergence of the second nature looks mysterious and up to a point *ad hoc*: "If we take law-governedness [i.e. the kind of understanding which is proper to empirical sciences] and spontaneity [space of reasons] to be mutually exclusive, it is difficult to see how lawful goings-on may be transformed into something which is subsequently no longer law-governed" (Gubeljic et al. 2000: 46). This is a familiar anti-normativist objection which we could call the "transition problem" (Turner 2010: 19-20). As Stephen Turner (2010: 2) stresses, the normativists should answer the question whether there is some kind of transformation from causal to normative, and then they should tell a story about the transition from one state to the other. The question could also be formulated like this: In what sense is second nature *nature*? What unifies first and second nature and how are they connected (Halbig 2008)?

4.1. McDowell's reply to the transition problem

McDowell's (2000: 97) reply is twofold. On the one hand, he stresses that second nature is the actualization of the potentialities that already belong to human beings' first nature. *Bildung* only actualizes the potential abilities that are part of the normal human organism. "[…] The innate endowment [i.e. the first nature] of human beings must put limits on the shaping of second nature that are possible for them" (McDowell 1998a: 190). In this sense, first and second nature are parts of a 'larger nature', for the latter cannot transgress the limitations of the former. On the other hand, he points out that the notion of second nature serves only the purpose of rejecting the idea that rationality is a supernatural power; it is not part of an evolutionary story about emergent

view of the relation between reason and freedom: rational necessitation is not just compatible with freedom but constitutive of it. In a slogan, the space of reasons is the realm of freedom" (MW: 5). Kant, although, lacked a "pregnant notion of second nature and this kept him from forming a completely right conception of experience even if he 'c[a]me so close" (MW: 97).

qualities. Thus, what we "need [is] only the bare invocation of *Bildung* not […] a detailed story about how what happens in *Bildung* connects with phenomena characterisable in terms of conformity to natural law" (McDowell 2000: 99). The aim of the argument is not to provide a straight answer to the transformation problem. McDowell (2008: 220) claims "that the only unity [he] need[s], to answer the question why the first-natural and the second-natural are both modes of the natural, is captured by the contrast with the supernatural, the spooky, the occult".

McDowell's argumentation takes a distinctively quietist course when he is asked to give a detailed reply on the relation between first and second nature,[13] for he declines to provide a positive or constructive answer. His sole concern is to show that no philosophical problem arises if we "accept that a distinctively human life is characterized by a freedom that exempts its distinctive phenomena from natural-scientific intelligibility, without thereby being required to push it back into the region of darkness, the region supposedly occupied by phenomena that resist the light cast by natural science because they are occult or supernatural" (McDowell 2008: 217).

4.2. Explanatory levels in first nature: a reply to the anti-normativist argument from the transition problem

Regardless of McDowell's quietism, I would like to suggest that his argumentation makes clear at least one thing: the so-called "transition problem" is a problem only as long as the distinction between first and second nature is interpreted in traditional ontological terms. However, the distinction between first and second nature does not aim to install a separation between two different ontological territories. On the contrary, it aims to dissolve every putative ontological schism. The distinction concerns two different "modes of intelligibility" (McDowell 2009: 262), namely two different ways of making things intelligible in nature (MW: 78, fn 8).[14] It has nothing to do with a claim about

[13] Note, for example, the following quote: "I do not need a knockdown argument that spontaneity is sui generis. All I need is to show how taking spontaneity to be sui generis does not pose the philosophical threats it can seem to pose" (McDowell 2002: 270)

[14] Elsewhere he defines the difference in the styles of explanation. He suggests that propositional attitudes "figure in a kind of explanation that is *sui generis*" in contrast to a pattern of explanation in which we make things intelligible "by representing their coming into being as a particular instance of how things generally tend to happen" (McDowell 1998b: 332 and 328 respectively). Other times, though, he expresses himself in a way that may give the wrong impression that the distinction between first and second nature has an ontological dimension: "I am quite happy to suppose there are two kinds of happenings in nature: those that are subsumable under natural law, and those that are not subsumable under natural law, because freedom is operative in them" (McDowell 2006: 238). I suggest we should resist this wrong impression exactly for the reasons I am unfolding in the present section.

the emergence of some ontologically distinct qualities. Thus, the only question that remains to be answered is whether the mode of intelligibility (or the mode of explanation) which is proper to second nature is a genuine one. The answer to this question does not necessarily involve any sort of evolutionary story concerning the transition from the non-normative to the normative state.

One way to elucidate this thought is to leave aside, for a moment, the second nature and the normative facts and take a closer look at the realm of first nature. In this realm, sciences make things intelligible by employing significantly different modes of explanation so that the same phenomenon can be cognitively captured in different scientific ways. For instance, physics, chemistry, and biology can treat the same phenomenon in different ways by explaining different aspects of it. A single human brain, for example, could be a system of electrons, protons, and neutrons according to a particle physicist; a system of chemical compounds that contains carbon (organic compounds) according to a chemist; a system of neurons for a biologist, or the centre of the nervous system for an anatomist. Each specialist cuts off a specific aspect of the same piece of reality and treats it by employing the conceptual tools proper to their domain.[15] What we get as a result are different answers to different questions which are based on significantly different patterns of explanation. In short, "Science explains the world at more than one 'level'" (Gibbard 2012: 2). In these cases, no one demands that, for instance, the biologist first tell an evolutionary story about how something biological emerges from something non-biological and then provide a biological explanation of a phenomenon. The problem for the biologist is whether her conceptual system can adequately capture the phenomena she studies, which is true of all the other disciplines as well.

Thus, the transition problem, if someone takes it to be a problem, may also concern the different scientific perspectives within the domain of first nature. It is not exclusively a problem of the emergence of the normative. The transition becomes problematic only from the standpoint which takes for granted that different modes of intelligibility should correspond to different ontological territories. This standpoint takes for granted that the existence of two kinds of intelligibility should correspond to two distinct ontological territories of

[15] One could object that I take for granted two ideas that I shouldn't without further ado: a) that all these scientists refer to the one and same reality and b) that the concepts of chemistry, biology and anatomy cannot be reduced to the conceptual system of physics. I am not ignorant of neither the radical idealist position that would reject the first idea nor the reductionist-physicalist one that would reject the second. However, this is not the place to argue against these two extremes. Therefore, I will take for granted the minimal realist position (i.e. the above-mentioned scientists refer to the same reality) and the minimal anti-reductionist position (the conceptual system of physics is not the only conceptual system that can cognitively capture objective reality).

the cosmos. Then the misleading question is raised: How does one produce the other? In this case, how does second nature emerge from the first? But the point of the argument for second nature is to exorcise the idea that different modes of intelligibility correspond to different ontological realms. Just as in the example above, the different ways available for scientifically treating a phenomenon do not entail the ontological diversity of this very phenomenon. There is no need to assume that, for example, the descriptions of physics and the descriptions of biology correspond to different ontological territories in order to accept that physics and biology employ different modes of explanation. The diversity of the explanatory patterns of the sciences can teach us that different modes of intelligibility can unproblematically be applied to the same ontological region. In fact, the anti-normativist demand for giving an account of the transition implicitly entails that there can be only one kind of intelligibility. If one *should* explain the transition from first to second nature that would mean that one *could* make both first and second nature intelligible by employing solely one explanatory pattern. This thought presupposes one ultimate way of making things intelligible. But in this case the anti-normativist begs the question. For as I already said, the argument in favor of second nature is an argument that suggests that there can be more than one kind of intelligibility of nature. Therefore, McDowell is correct in saying that he doesn't need "a detailed story about how what happens in *Bildung* connects with phenomena characterisable in terms of conformity to natural law".

McDowell's response to the anti-normativist argument from the "transition problem" makes clear a few things about the distinction between first and second nature: 1) The distinction is not ontological at all. It refers to two different ways of making things in nature intelligible, that is, two different kinds of explanation (normative and empirical-scientific). 2) Both kinds of intelligibility refer to one ontological territory, that is, nature. 3) There is no philosophical need for presenting a philosophical evolutionary story in order to explicate how phenomena explained normatively emerged from phenomena explained scientifically. I attempted to make this position stronger by examining the relation between different levels of explanation within the general framework of empirical-scientific understanding.

5. *Science-informed philosophy*

The questions which arise now are these: can the law-governedness be the essential feature of the mode of intelligibility which is proper to first nature and contrasts the kind of intelligibility which is proper to the normative facts? Is second nature the necessary and sufficient condition of the kind of intelli-

gibility that is proper to reason? As we will see, McDowell's post-*MW* reply is negative to both questions.

In *MW* (109) McDowell suggests that the scientific revolution of the 17th century brought about the conception of nature as the realm of law, which was not available in the ancient and medieval era. Furthermore, he stresses that defending the genuineness of the kind of intelligibility which is proper to the space of reasons does not imply a rejection of the conception of nature as the realm of law. In short, he attempts to adjust his philosophical perspective to the evolution of the sciences and especially to the conceptual breakthrough that took place during the scientific revolution. His conception of nature takes into serious consideration the developments of empirical sciences and in this sense, he provides a science-informed philosophical outlook.

5.1. Second nature and biology

The fact that McDowell's philosophy is or attempts to be science-informed is shown also in his later writings where he acknowledges that in *MW* he works "with an unsatisfactorily monolithic conception of what is to be contrasted with the distinctive kind of intelligibility for which responsiveness to reasons constitutes the framework" (McDowell 2000: 98). The main problem with this monolithic Russellian[16] conception is that it suits only to a small part of the natural sciences, that is, mathematical physics. It does not do justice to the explanatory patterns of other sciences and especially biology which is more relevant to functions and processes that are related to human beings (McDowell 2008: 220). Furthermore, when the discussion comes to biology it is revealed that non-human biological creatures – like a trained dog – can and do have a second nature (McDowell 2000: 99; 2006: 236; 2008: 220). Thus, the set of all second-natural phenomena is not coextensive with the set of phenomena that should be made intelligible by employing the kind of intelligibility which is proper to the space of reasons. For instance, the phenomena of Pavlovian conditioning are second-natural phenomena (related both to humans and to other animals) but they do not require an intelligibility of the space of reasons variety. In short, what needs to be revised in *MW*-conception of nature is a) the idea that first nature can be equated to the realm of law; and b) that all second-natural phenomena could be understood by employing the kind of intelligibility which is proper to reason.

[16] This conception about scientific explanation is central to the so-called 'received view' in the philosophy of science and it finds its paradigmatic form in Hempel's and Oppenheim's (1948) Deductive-Nomological model. For a historical presentation on the issue of scientific explanation, see Cartwright (2004).

McDowell revises his account by taking into consideration a richer conception of natural sciences which does not include solely mathematical physics and the explanation that is based on law-governedness. Bringing biology into his philosophical picture helps him realize that second nature is not a distinctive feature of human beings and that second nature is only a necessary, but not a sufficient, condition for rationality. The acquisition of rationality presupposes the second nature, but the second-natural phenomena does not necessarily entail the existence of rationality. There are second-natural phenomena that require a kind of intelligibility which is not substantially different from the intelligibility required to capture the first-natural phenomena (McDowell 2008: 220). Only a subset of the second-natural phenomena is made intelligible by the placement to the space of reasons.

This revision makes McDowell's account more refined and even more science-informed. It incorporates a richer and more accurate conception of the explanatory patterns that the natural sciences employ, but it does not change its principal thought: the kind of intelligibility that is a matter of placement in the space of reasons is *sui generis*, that is, "beyond the reach of the natural-scientific understanding" (McDowell 2008: 217). And still it is concerned with phenomena that belong to nature: they are not spooky, occult or supernatural. In the *MW*-version of the account these phenomena were identified with the second-natural phenomena whereas in the later version they are only are identified with a subset of the second-natural phenomena.

5.2. The naturalistic threat of human sciences

So far, so good--it seems that taking lessons from the sciences does not pose any threat to the liberal naturalist account of McDowell. However, I would like to argue that McDowell's account stops taking into consideration the lessons from the empirical sciences exactly at the point where the threat is less acute, although it is not entitled to stop at this 'convenient' point. My question is this: what happens if after biology we try to insert human sciences into our philosophical image? Biology clearly[17] leaves room in the set of second-natural phenomena for a kind of intelligibility that is not proper to the biological explanatory patterns. If we add human sciences in our picture though, no such room is left. The subject matter of the human sciences covers the entire range of human actions and beliefs. Besides the part of the second-natural phenomena that are cognitively captured by biology,[18] no other section of second nature

[17] Clearly for everyone who is not some kind of radical biological reductionist.

[18] These by definition don't belong to what can be made intelligible by placement in the space of reasons.

can – at least in principle – be ruled out by the reach of human-scientific understanding. It seems that the emergence of human sciences brings with them a much more acute philosophical anxiety than this one of natural sciences in directing straight forward to threaten the autonomy of the space of reasons, by trying to make human's second nature fully intelligible through modes of explanation which are alien to the placement in the space of reasons.[19] Bald or neo-Humean (McDowell 1998) naturalists could argue that the scientific revolution and the corresponding emergence of mathematical physics saved us from the ancient and medieval superstition that first nature is the realm of final ends; furthermore, the emergence of biology and modern medicine exempted a part of the second nature from the kind of intelligibility which is proper to reason; and now the human sciences exempted the rest of second nature from this peculiar or supposedly *sui generis* kind of understanding. The more we examine reality the more we realize that the only legitimate kind of intelligibility is the perspective of empirical science. After all, the scientistic naturalist could conclude, the only science-informed philosophy is this kind of naturalism which equates nature with the subject matter of the empirical sciences.

In short, if the notion of second nature is not beyond the reach of empirical-scientific explanations, what is the argument against the bald naturalization of every single aspect of human acting and thinking? Second nature as introduced by McDowell, in order to demarcate the phenomena which should be made intelligible by placing them into the space of reasons. If part of the second-natural phenomena can be reached by biological understanding, what prevents us from thinking that the rest of them can be reached by human-scientific understanding?

6. *Normativist strategies for dealing with human sciences*

In general, the normativist strategies against the bald naturalist threat posed by the emergence of human sciences can take three different routes. The first choice is to argue that human science does not exclude normative explanations and hence they do not employ the same explanatory patterns as the natural sciences. I will call this choice 'hermeneutic conception of human sciences'. The second choice is to provide a demarcation criterion for singling out the cases which should be made intelligible by empirical-scientific understanding and the cases which should be made intelligible by placing them in the space of reasons. I will call this choice 'traditional demarcationism'. The third choice is to argue that while there is no pre-existing criterion for demarcating cases of

[19] By placing things into the "ordinary stream of explanation" (Turner 2010: 11).

genuine expression of rationality, the kind of intelligibility which is proper to reason is ineliminable. I will call this choice 'normative fallibilism'. In the rest of the present section I will attempt to show why the first two options are not viable and why we should endorse the third.

6.1. The hermeneutic conception of human science

One way to neutralize the bald naturalist threat that comes with the human sciences is to argue that the latter employ by definition the kind of intelligibility that is proper to reason. This line of thought has its roots in the Hermeneutic tradition of the 19th century German-speaking world which took a mature form in Wilhelm Dilthey's work (Harrington 2001: 43) but is also central to a particular trend in contemporary philosophy of social sciences (see Risjord 1998: 224, Babich 2017). Hermeneutics imply a division between *Naturwissenschaften* (sciences of nature) and *Geisteswissenschaften* (sciences of spirit) not only on the level of the scientific object but also on the level of the characteristic modes of explanation. The difference is often expressed by using another two German terms: Sciences of nature "involve *Erklären* (explanation by way of laws) while [sciences of spirit] involve *Verstehen* (hermeneutic understanding from the "inside")" (Macarthur 2010: 134). This means that human sciences make things intelligible by employing some kind of "empathetic understanding" (Stueber 2012), that is, by taking into consideration the reasons that brought about a specific situation. This is the *sui generis* style of understanding things in the meaningful territory of human actions and beliefs. Thus, according to this conception, the emergence of human sciences does not pose a naturalist threat at all. Normative explanations are found at the very heart of human-scientific understanding and hence the latter is not a threat for the former.

However, this choice suffers two major problems. First, the evolution of the human and social sciences does not seem to vindicate this sharp distinction between natural and human sciences. The unity of natural and human sciences can be revealed by appealing to the differences within the two categories: "[t]here is as much reason to think that there are significant differences *within* the category of natural science, and *within* the category of human science, as there are *between* the natural and the human sciences. This vitally important point is gradually gaining credence" (Macarthur 2010: 134, emphasis in original). The rejection of the monolithic nomological model of natural-scientific explanation created a wider conception about scientific explanations. Within this wider conception there is no sharp distinction between the explanatory patterns employed by the natural and the human sciences. From the standpoint of this wider conception we can point out, for instance, that the "[t]he types

of knowledge gained of the social world are much like the types of knowledge we can claim of the biological world" (Mitchell 2009: 131). Second and most important for my argumentation, the hermeneutic choice does not avoid the philosophical problem of the relation between nature and reason, it just pushes it into another philosophical territory. Instead of explaining why normative explanations are genuine and cannot be altogether reduced to the empirical explanations provided by the human sciences, the normativist has to explain why normative explanations within the domain of human sciences are genuine and cannot be reduced to explanations of another type. The philosophical need for providing an argument against the eliminability of normative explanations remains.

6.2. Traditional demarcationism

The second option accepts that human sciences do not provide normative explanations and that they may pose a naturalistic threat for the philosophical image of our intellectual and practical freedom. However, this option presupposes that the genuine expression of rationality cannot be made intelligible by empirical sciences and proposes a criterion (or a set of criteria) for demarcating between cases of genuine expression of rationality and other cases. This option is characteristic of a Kantian line of thought. This is, for instance, the position that Sebastian Rödl (2007) and Christine Korsgaard (1996; 2009) have recently defended.

Rödl follows Kant in considering that the idea of freedom is the idea of a certain kind of determination, namely a certain kind of causality: a causality of thought. He also follows the principal thought of German idealism that reason, self-consciousness, and freedom are one (Rödl 2007: 105). Thus, expressions of freedom are also expressions of rationality and can be understood only by employing the kind of intelligibility which is proper to reason. "Being free is being subject to a causality of thought and, hence, is placing oneself under an order of reason" (Rödl 2007: 112). As Kant teaches, free will is autonomous in being subject to the laws that are its own, which are the laws of reason. According to Rödl, this does not mean, as some contemporary interpretations[20] of Kant suggest, that free will legislates lawlessly the laws that are its own, which would be a paradox. It means that 'one's own' [does] not signify the origin of the law. It […] signify its logical form, the kind of law that it is' (Rödl 2007: 117). The logical form of a law of autonomy is determined in distinction to the logical form of a law of heteronomy: while "a law of autonomy explains acts that exemplify it by the nature of the subject of this act and by it alone"

[20] See Pinkard (2002) and Pippin (2008).

(Rödl 2007: 119), "a law of heteronomy is one according to which one thing is determined to act by another thing" (Rödl 2007: 118). Thus, in this view, we can demarcate the cases of genuine expression of rationality by appealing to the logical form of the law that explains what is going on in every case.

Korsgaard proceeds in a different but analogous way, appealing directly to the Kantian imperatives (both hypothetical and categorical) as the constitutive principles of an intentional action (Korsgaard 2009: 92), that is, as the constitutive principles of the expression of rationality at the level of action. According to Korsgaard, the reflective structure of the human mind and the consequent "reflective distance from our impulses makes it both possible and necessary to decide which ones we will act on: it forces us to act for reasons" (Korsgaard 1996: 113). When we succeed in acting according to the categorical imperative, we can consider ourselves autonomous and placed in the realm of normativity. The categorical imperative governs our (free) actions through the mediation of the various practical identities we endorse (Korsgaard 1996: 101). The details of Korsgaard's account are not the point here. The point is the general philosophical aspiration which is common to Rödl's account and which seeks the demarcation between cases of genuine expression of rationality and cases of mere causal events.

This demarcatonist option though is vulnerable to the sceptical anti-normativist arguments. Anti-normativists like Turner take into consideration the diversity of normative contexts in the world and in history in order to cast doubt on the genuineness of the normative explanations. The argument goes as follows: since "most of the people in history and in the present were and are living in normative error" (Turner 2010: 181), the appeal to some binding rules of reason is at stake. The argument shares the structure of the pessimistic meta-induction against scientific realism. In this well-known argument it is recalled that most of the scientific theories that were once successful in the past are now considered false: "Therefore, the pessimist concludes, current successful theories will turn out to be false as well" (Mizrahi 2013: 3210). Correspondingly, Turner concludes that the appeal to some genuine normative force that obliges human beings and the relevant demarcation of the cases of authentic expression of rationality is chimerical, since most people have lived in normative error. The error is revealed since we are in a position to explain people's actions and beliefs not by appealing to the truth or goodness of their beliefs and actions, respectively, but by appealing to various biological, psychological, or sociological causal factors. People in history change their minds about what counts as true or good, and this jeopardizes the normativist effort to demarcate between genuine expressions of rationality and mere causal events. In short, since every proposed normative explanation has been reduced

to various empirical-scientific explanations we have good reasons to believe that also in the future the normative explanations are going to be reduced to scientific explanations. Therefore, as the history of human thought shows, no demarcation criterion is able to single out the cases of genuine expression of rationality and consequently to secure the ineliminability of the domain of normative explanations.

6.3. Normative fallibilism

Despite its plausibility, we have to note that the sceptical argument from the normative error threatens only some versions of normativism, and not normativism *per se*. The argument is valid only against the versions of normativism which suggest that the genuine expressions of rationality can be demarcated infallibly. Other versions of normativism, like McDowell's liberal naturalism, are not threatened by the argument from normative error. In contrast to a Kantian line of thought, McDowell does not seek for a demarcation criterion for singling out the cases of genuine expression of rationality, i.e. the genuine cases of autonomy. In this sense, his account is not vulnerable to the skeptical argument from normative error. He stresses that while autonomy is a capacity[21] which can fail, this does not prove that it doesn't exist at all. Sometimes a rational subject acts or thinks not by responding to genuine reasons but only to what seems to it to be a reason. This point is well taken by the anti-normativist argument from the normative error. And, of course, this entails that we need to distinguish between two kinds of facts: a) the genuinely normative facts and b) the seemingly normative facts. But on this issue "there is no criterion, if by that we mean some general formula that it might be possible to apply to mark off genuine reasons from impostors", on the contrary

> the only thing one can do is to ask oneself whether it coheres with one's view of other regions of the space of reasons, which one must simply trust for the duration of one's reflection about the region that is under reflective scrutiny. In a familiar image, one's possibilities for reflection are those of Neurath's mariner, repairing his vessel, or at least inspecting it for seaworthiness, while it is afloat. One can be confident only that if one has the space of reasons in view at all one cannot be completely wrong about it' (McDowell 2010: 12).

Even if we don't agree with McDowell that this is "the only thing one can do", we can agree that it is not necessary to be committed to the formulation of an infallible criterion for demarcating the genuinely from the seemingly

[21] The capacity "to subject oneself to the normative force of reasons" (McDowell 2010: 9).

normative facts in order to accept that genuinely normative facts exist. The sceptical argument from the normative error is only effective against the normativist accounts which aim to provide a demarcation criterion for singling out the cases of authentic expression of rationality.

Normative foundationalism (i.e. the determination of an infallible criterion for demarcating the cases that need to be explained normatively) and normative scepticism (i.e. the rejection of the genuineness of the normative explanation on the basis of their fallibility) are not our only options. We can argue that human beings do possess the capacity to respond to reasons and therefore to determine themselves, even if sometimes they fail to do that. Hence, even if we acknowledge that each and every normative explanation can be reduced to an empirical-scientific explanation, we can reject the idea that the domain of the normative explanations is altogether eliminable. Another way to express the middle way between normative foundationalism and normative scepticism is to say "that the epistemology of values or duties is Neurathian, meaning that there is no one-way, axiomatic structure to such normative reasonings" (Blackburn 2001: 150). This is the perspective I call 'normative fallibilism' and for the reasons I attempted to show is the only viable strategy for defending a normativist perspective.

7. *Reduction and Eliminability: a constructive reply to scientific naturalism.*

Let me now recapitulate the philosophical image that I have already sketched by following some central tenets of McDowell's liberal naturalism. 1) The notion of second nature leaves room for normative explanations without appealing to any kind of supernatural entities, cognitive powers, or phenomena. 2) Second-natural phenomena are not co-extensive with phenomena explained normatively, for second-natural phenomena also includes phenomena that need to be explained by biology and human sciences. 3) There is no *a priori* criterion for demarcating second-natural phenomena which should be explained normatively and second-natural phenomena which should be made intelligible by empirical-scientific understanding.

At this point the scientific naturalist can insist: maybe normative fallibilism is a less vulnerable position than traditional demarcationism but it does not provide any argument in favor of the idea that some phenomena in nature rest outside the reach of the empirical-scientific understanding and hence that they should be made intelligible by exclusively employing normative explanations. Given the argumentation I have already provided, we cannot defend the genuineness of normative explanations by appealing solely to the notion of second

nature and we also cannot do that by proposing a demarcation criterion for singling out the case of genuine expression of rationality. If we further rule out the obsolete choice of the hermeneutic division between *Naturwissenschaften* and *Geisteswissenschaften*, then our defence against scientific naturalism looks weak. McDowell's quietist strategy consists in repeating that no philosophical worries arise if we place rationality into the realm of nature and that the naturalism of natural science should not be taken as a default position. The default view should be that "human beings are unique among living things – *outside the reach* of the sort of understanding achievable by a scientific biology – in virtue of the freedom that belongs with our responsiveness to reasons as such, […] unless it can be shown to be wrong" (McDowell 2006: 237, emphasis is mine). But taking into account the human sciences, I think, makes McDowell's "naturalism of second nature" more vulnerable to the bald naturalistic threats and his quietist strategy less convincing. I don't see why it is self-evident that human beings are unique among living things. In this sense, I can't see how it can be convincing that there is something outside the reach of understanding which is proper to the empirical sciences.

7.1. Reducibility and eliminability

It is exactly at this point where my argumentation diverges gravely from McDowell's defence of normativism, for in my view there is need for a constructive (as opposed to quietist) argument against the naturalist threat of eliminating the kind of intelligibility which is proper to reason. I think that this argument can be found in distinguishing between the concept of explanatory reducibility and the concept of the eliminability of the domain of normative explanations. I suggest that scientific naturalism seems plausible only because we think that the former notion necessarily entails the latter and I want to provide an argument against this idea.

Let me start by giving an example. Let's suppose that someone, say X, believes in creationism and also, she does the housekeeping. Thus, 'X believes in creationism' is an example of a belief and 'X does the housekeeping' is an example of an intentional action, both of which need to be explained. If X is asked why she believes in creationism and why she does the housekeeping, she can give a normative explanation of both the belief and the action. She can say, for instance, that the theory of evolution has many gaps and thus it is more plausible to think that we have been created by God. Also, she can say that she does the housekeeping because she is a woman and this is what women do. Women do the housekeeping and men provide financially for the family. Those are examples of normative explanations of the above-mentioned belief and intentional action. Believing in creationism is explained by conforming

to the epistemic norms of consistency and theoretical completeness while doing the housekeeping is explained by conforming to the practical norm of the gender-based division of labor. Consider now the case where the normative explanations in question are reduced to empirical-scientific explanations. For instance, a sociologist can assert that X believes in creationism because she grew up in a community which attempts to preserve its cohesion through religion. Correspondingly, a psychologist may assert that X has this gender-based behavior because of her dominating father figure. These are the paradigmatic cases for scientific naturalists. They argue that since the human sciences have managed to reduce some of the normative explanations to empirical-scientific explanations there is nothing that cannot be reduced to the empirical-scientific understanding. In other words, they assume that since at least some of the normative explanations have been proved mistaken and reduced to scientific explanations, then gradually all normative explanations are going to be reduced to scientific explanations. Thus, they assume that the reducibility of normative explanations entails the eliminability of normative vocabulary. But I think that this entailment is mistaken.

First of all, in order to reduce the normative explanations to the scientific explanations we need the normative vocabulary at our disposal. Reducing X's normative explanations about believing in creationism to the sociological explanation about her community we have to think that it is correct that her community sought for cohesion through religion, and that this situation caused her beliefin creationism. In an analogous way, reducing X's normative explanation about gender roles to the psychological explanation involving her father figure we have to think that it is true that the dominating father figure is causally connected with the sexist conception about the division of labour. But thinking that the psychological or the sociological explanations are true presupposes that they conform to some kind of epistemic norms. Therefore, the idea that we can eliminate the normative vocabulary is incoherent. The point here is that the reduction of the explanations which are proper to the space of reasons to something alien to them is always the work of reason, for each instantiation of reduction also consists of justifying and being able to justify this very act of reduction. Thus, the image of the reduction of the normative explanations which gradually leads to the elimination of the normative vocabulary is misleading.

7.2. God's point of view

What makes this image plausible relies on the presumption that the cognitive act of the reduction is undertaken from a standpoint external to the bearer of actions or beliefs of which normative explanations have been reduced to scientific ones. In one sense, it presupposes that the object of reduction is neces-

sarily different from the subject of reduction. In our examples, it presupposes that X never embraces the content of the explanatory reduction. But again, this presumption is misleading. Of course, sometimes this turns out to be the case. For instance, I can take advantage of my knowledge of empirical sociology and psychology and arrive at the conclusion that X's beliefs and actions are the result of a specific social milieu and a particular psychological condition. And maybe X never draws this conclusion herself. In this case, I don't have to take into account X's space of reasons, and consequently X's space of reasons vanishes from my picture. However, whether or not the actual person X has access to the content of the scientific explanation is a mere contingency which has no philosophical interest. What matters from a philosophical point of view is that the act of reduction is at X's disposal. This means that the paradigmatic case for our philosophical conception of the issue should not be when someone other than X realizes that X believes or does something because of some sociological or psychological factors but when X herself realizes that. In this latter case, X's space of reasons does not disappear from our philosophical picture, it is just modified. After the realization (i.e. after embracing the content of reduction), X will still adopt a normative attitude about both the roots of human life and the division of labor, and she will do so by subsuming herself to the dictates of her reason. The difference is that now her space of reasons is modified. And it is not simply modified; it is *enriched* in an important sense. If the acts of reduction are correct, X's new space of reasons contains a few more justifications that prevent her from making a mistake. During her reflective scrutiny about what is a reason for believing something or acting in a particular way, taking into consideration the sociological or the psychological knowledge about her condition can prevent X from believing uncritically that we are God's creatures or that women are supposed to do the housekeeping.

Therefore, the reduction of a normative explanation to a scientific explanation is not an episode toward the gradual shrinking of the space of reasons, but an episode toward its expansion. Equating reduction with elimination presupposes that the subject of reduction is necessarily different from the object of reduction, and this entails that there is something like a super-reason that accomplishes the task of reduction, a super-reason that is not and could not be affected by this very cognitive act of reduction. Envisaging the space of reasons as shrinking after every act of reduction requires the putative standpoint of a reason that is not affected by this cognitive act. However, this standpoint would be something like God's point of view: an external vantage point overarching thought and world. Only from this standpoint can the normative vocabulary be considered eliminable. But the presupposition of this sort of vantage point is very problematic, especially for accounts which aim to be naturalistic.

7.3. Genuine normative explanations and the scope of empirical-scientific understanding

Hence, in one very essential sense, scientific explanations are constitutive for our freedom rather than a threat to it. They help us become freer. They prevent us from normative error and consequently from being mere slaves of the various causal goings-on. Scientific explanations can decisively contribute to the reflective scrutiny which determines what is a genuine reason for a belief or action and what is merely an impostor. In each case empirical-scientific understanding can be part of the process of determining what is a genuine (as opposed to seemingly) normative explanation. We can consider as genuine the explanations which resist their plausible reduction to empirical-scientific understanding. Of course, our estimation of which normative explanations resist such a reduction is clearly fallible and depends on various historical factors (the maturity of empirical-scientific understanding, the political institution of our society, etc.). But the concept of resistance to the reduction to empirical-scientific explanations is the only concept we need in order to arrive at the concept of the genuine normative explanation.

By saying that the concept of the genuine normative explanation needs only the concept of resistance to the reduction to the empirical-scientific understanding I imply that the concept of the genuine normative explanation does not presuppose any sort of "super-added normative element that cannot be accounted for naturalistically or by social science", nor does it entail that the genuineness "of the reasons, have some sort of explanatory force beyond the mere 'natural' fact of people's beliefs and desires" (Turner 2016: 10 and 9 respectively). We can admit that nothing is beyond, outside, or above the reach of empirical-scientific understanding without endorsing any version of anti-normativism. When we say that a phenomenon is genuinely normatively explained we do not have to mean that the phenomenon is beyond the reach of empirical-scientific understanding. Take, for instance, the case that I believe that there is a computer in front of me. I can normatively explain the belief 'there is a computer in front me' by showing how it is conformed to an epistemic norm, say the norm that dictates to 'endorse the impression you perceive under normal lighting conditions'. This phenomenon is not beyond the reach of biology, psychology, sociology, or physics. Several things can be said about light's wavelength, my retina, my psychological condition, or the social milieu of my upbringing. All are certainly connected with my belief that there is a computer in front of me. But if I have followed the epistemic rules correctly and there is indeed a computer in front of me then my belief that there is a computer in front of me cannot be explained without showing how this belief conforms to some epistemic rules. That is a genuine normative explanation.

If this is the case, we don't need to adopt the idea that there is a peculiar super-added normative element in the phenomena which needs to be explained normatively, and we don't need to admit that there is something beyond the reach of empirical-scientific understanding in order to defend normativism. The conceptual distinction between the reducibility and the eliminability of the normative explanations may prevent us from seeking something that is outside the reach of scientific understanding. I think that this sort of talk is misleading and even suspect of supernaturalism. Something that is beyond empirical-scientific understanding as such, and by that we should include both present and future[22] sciences, looks horrifyingly indistinguishable from something that is beyond nature itself. By disassociating the concept of explanatory reducibility from the concept of the eliminability of normativity there is no need to invoke something beyond the reach of empirical-scientific understanding. We can appeal to just another equally legitimate way of making a phenomenon intelligible. Maybe even the use of the word 'space' in the famous Sellarsian terminology of 'space of reason' is not the more appropriate term, for it creates a picture of a bounded territory which is exactly outside or beyond another territory which is occupied by empirical-scientific understanding. Maybe the vocabulary of relation could be more instructive here. Freedom consists in a special relation of rational subjects with themselves and with others. The more those subjects realize the causal goings-on that govern these relations the more the subjects find themselves in the normative realm of freedom.[23] This terminology can avoid the misleading spatial metaphor about normativity.

8. *Integrating first and second nature*

My main argument is that by rejecting the idea that explanatory reducibility entails the eliminability of the domain of normative explanations, we can create the suitable logical space between scientific naturalism and supernaturalism. The interchange between the two modes of intelligibility – one proper to reason and the other proper to mere causal events – protect us from either depicting our rationality as an occult power or presenting it as

[22] This aspect of the problem has been revealed by the famous Hempel's dilemma. See a variation of it which is closer to my discussion in Macdonald 2008.

[23] This sort of vocabulary is absolutely related to German idealism and especially to Hegel. As Pippin (2008: 4) stresses, Hegel's perspective "has two basic components: that for Hegel freedom consists in being in a certain reflective and deliberative relation to oneself (which he describes as being able to give my inclinations and incentives a "rational form"), which itself is possible, so it is argued, only if one is also already in certain (ultimately institutional, norm-governed) relations to others, if one is a participant in certain practices".

an illusion, and leads us to a proper liberal naturalist perspective. The perspective I am proposing is naturalistic insofar as it leaves nothing 'beyond the reach' of scientific understanding and presents scientific explanations as constitutive of the space of reasons. It is also liberal in the sense that it rejects the eliminability of normative vocabulary. Rationality is not a mysterious power outside of nature but a capacity to take control of our lives by understanding how the causal goings-on work, that is, by gradually knowing more about the cases in which we do not have control of our lives. Furthermore, it is impossible to get rid of the kind of intelligibility which is proper to reason without appealing to the extremely questionable idea of God's point of view. Therefore, we can say that the essential feature of second nature – or more correctly of the part of second nature that is made intelligible normatively – is human ability to grasp cognitively the causal goings-on that is part of both first and second nature. In this sense, there is absolutely no philosophical gap between first and second nature and there is absolutely no philosophical need to explain how causal goings-on may be transformed into something which is subsequently no longer causally ordered. Second nature is the purely natural ability to adjust our beliefs and actions to the knowledge of the causal-goings on that dictates our lives.

The perspective I have tried to defend has two main philosophical consequences that are not entailed, at least immediately, by McDowell's liberal naturalist account. The first is that freedom is a concept that allows for degrees (Pippin 2008). The second is that the layout of the space of reasons is historically changeable. In what follows, I can give only a sketchy account of these two consequences.

8.1. The degrees of freedom

If the expressions of our freedom are identified with the expressions of our rationality and if the expressions of rationality are proportional to our knowledge of the causal order that dictates our lives, it follows that we can only determine various degrees of freedom instead of absolute states of freedom. These various degrees of freedom depend on our knowledge of the causal goings-on that affect us. The more we know the various physical, biological, psychological, or sociological factors that affect our lives the more we can endorse beliefs and undertake actions of which endorsement or undertaking cannot be explained by appealing to empirical-scientific explanation and hence, can be considered as an expression of our rationality and consequently of our freedom. This is largely the philosophical 'cost' we have to pay for rejecting the idea that we can formulate an infallible criterion for demarcating the cases of genuine expressions of rationality and adopting

normative fallibilism. Only if such a criterion were feasible could we refer to the expression of rationality and hence to our freedom as an "'either you have it or you don't' capacity" (Pippin 2018: 214).

8.2. Reason's historicity

The second consequence of the conception I have attempted to defend, which is closely related to the first, is that the content of reason or in McDowell's (2018) terms the layout of the space of reasons is historically changeable. If I am correct, an indispensable part of the reflective scrutiny which determines what is a genuine reason (as opposed to an impostor) is the examination of what normative explanations can be reduced to empirical-scientific ones. The acts of reduction modify (enrich) our conception of what is a reason for what. In this sense, the empirical-scientific knowledge concerning the biological, psychological, or sociological factors that affect our beliefs and actions shapes drastically the layout of the space of reasons. These acts of reduction are historically determined, for they are subject to the maturity of the empirical-scientific understanding and to the various institutions that can make empirical-scientific knowledge available to people. Therefore, the layout of the space of reasons is also historically determined.

McDowell, in a recent debate with Pippin, rejects explicitly this idea both as a plausible interpretation of Hegel's philosophy and as truth. Commenting on a specific example, the rejection of the gender-based division of labor after the 1970's, Pippin (2018: 217) argues that it doesn't seem plausible that we rejected this kind of division just because we suddenly "had our eyes wider opened". Pippin's point is that we should not think that the moral truth about the wrongness of gender-based division of labor was always there and we just discovered it at some historical point. We have to think that the layout of the space of reasons has historically changed. On the contrary, McDowell (2018: 254) argues that it would be absurd to consider that the layout of the space of reasons has changed because this would mean that before the 1970's there was a moral basis in this kind of division of labor. "It [just] used to be thought that there was a moral basis". McDowell's point is that we can't think that something used to be morally (or epistemically, I could add) right and now is morally (or epistemically) wrong. We have to think that up to a point we used to believe that it is right and then we discovered that it is wrong.

I think that McDowell's view on this particular issue is mistaken, for he conflates the historical changeability of the layout of the space of reasons with the changeability of the truth-value or the moral-value of epistemic or moral judgements. In short, McDowell believes that if we accept that the layout of the space of reasons is changeable, we are led to the idea that the moral-value

or the truth-value of some judgements can historically change from right to wrong and *vice versa*. And he believes that this idea is incoherent. But I would like to argue that the idea of the historicity of reason has no such implications. It is not the case, for instance, that slavery was indeed morally right and then became morally wrong or that it used to be true that the aether exists but now it is not true. The reason we have to accept the historical changeability of the layout of the space of reasons is not that there can be a genuine change in the truth-value or the moral-value of some judgements. The reason is that some judgements, up to a historical point, are not even candidates for truth-value or moral-value.[24] In Pippin's and McDowell's example, we don't have to think that the gender-based division of labor was morally right before the 1970's and became morally wrong since then. We just have to think that the judgement concerning the gender-based division of labor, up to a historical point, was not even candidate for receiving moral value (right or wrong). It is only after this judgement became a candidate for receiving a moral-value that we can say that this kind of division was always wrong. Becoming a candidate for moral-value is the result of a specific historical course which, as I argued, includes the maturity of empirical-scientific understanding and the evolution of the social institutions. Thus, judging that gender-based division of labor was always morally wrong is possible only *ex post facto*, only after this particular judgement became a candidate for moral-value. But it is exactly on this ex-post-facto-ness that the historicity of reason relies.

In short, it should be not considered that the idea of the historical changeability of the layout of the space of reasons entails the idea that moral or epistemic judgements can genuinely change truth-value or moral value. It just entails the idea that the space of reasons is reorganized when new judgements become candidates for truth-value or moral-value. As I attempted to show, this is part of a historical process which includes the ongoing expansion of our empirical-scientific knowledge of the causal factors that dictate our lives.

9. *Conclusions*

McDowell's liberal naturalism provides the fundamental conceptual tools needed in order to make logical room between a philosophical stance which legitimizes only the kind of intelligibility which is proper to the empirical sciences and the philosophical stance which presents human rationality as

[24] As Joseph Rouse (2016: 30) stresses, "people can now tell and talk about mitochondria, the Precambrian Era, subatomic particles, tectonic plates, retroviruses, spiral galaxies, and chemical kinetics. One need not go back very far historically to find not error but silence on these and so many more scientific topics".

supernatural power. The most valuable of these tools is a) the disassociation of the idea of nature from the idea of the subject matter of the natural sciences and b) the rejection of the idea that there can be a criterion for demarcating the genuine from the seemingly normative facts (what I called normative fallibilism).

Based on McDowell's account I attempted to show that a viable liberal naturalist account should endorse three additional interconnected assumptions: a) the disassociation of the concept of the reducibility of the normative explanations from the concept of the eliminability of the domain of normative explanations, b) the untenability of God's point of view, and c) the rejection of the view that normativism presupposes the idea that there is a super-added normative element which is beyond the reach of empirical-scientific understanding. I also argued that the above-mentioned assumptions have two philosophical consequences that diverge from McDowell's point of view. The first is that freedom as the expression of rationality is a concept that allows degrees depending on the historical context. The second but related to the first is that the layout of the space of reasons is historically changeable.

Let me now provide one brief final remark. In the beginning of the present text I suggested that the account I propose attempts to understand normativity in a collectivistic and historical way in contrast to ahistorical and individualistic philosophical accounts. I think that the historical character of my approach should be already clear. On the contrary, the collectivist dimension may be not so obvious. Thus, I should stress that a presupposition of what I have already said is that the unit of my philosophical analysis is not the individual subject. If it were the individual subject we would continue to be trapped in the oscillation between a version of normativism which seeks for a space beyond the reach of scientific understanding and the scientistic rejection of this sort of space and of normativity altogether. If we take the problem of rationality to be a problem of whether or not there exists some sort of capacity at the level of individuals, then we can easily imagine an external standpoint. Given the external standpoint the distinction between reducibility and eliminability cannot be made. But if we focus at the level of tradition (historical community), the external standpoint vanishes and the distinction between reducibility and eliminability becomes possible.

Thodoris Dimitrakos
thdimitrakos@phs.uoa.gr
University of Athens, Hellenic Open University

References

Babich, B., ed., 2017, *Hermeneutic Philosophies of Social Science*, De Gruyter, Berlin.

Bilgrami, A., 2010, "The Wider Significance of Naturalism: A Genealogic Essay", in M. DeCaro and D. Macarthur (eds.), *Naturalism and Normativity*, Columbia University Press, New York: 23-54.

Blackburn, S., 2001, "Normativity à La Mode", in *The Journal of Ethics*, 5 (2): 139-153.

Broome, J., 2007, "Is Rationality Normative?", *Disputatio*, 2 (23): 161-178.

Chan, T. (ed)., 2013, *The Aim of Belief*, Oxford University Press, Oxford.

Cartwright, N., 2004, "From Causation to Explanation and Back", in B. Leiter (ed.), *The Future for Philosophy*, Oxford University Press, Oxford: 230-245.

Davidson, D., 2001, "Mental Events", in *Essays on Actions and Events.* Oxford: Clarendon Press: 207-224.

Finlay, S., 2010, "Recent Work on Normativity", in *Analysis*, 70 (2): 331-346,

Gibbard, A., 2012, *Meaning and Normativity*, Oxford University Press, Oxford.

Harrington, A., 2001, *Hermeneutic Dialogue and Social Science: A critique of Gadamer and Habermas*, Routledge, London-New York.

Halbig, C., 2008, "Varieties of Nature in Hegel and McDowell", in J. Lindgaard (ed.), *John McDowell: Experience, Norm, and Nature*, Blackwell, Singapore: 72-91.

Hegel, G.W.F., 1977, *Phenomenology of Spirit*, Eng. trans. A.V. Miller, Oxford University Press, Oxford.

Hempel, C. and Oppenheim P. (1948), "Studies in the Logic of Explanation", in *Philosophy of Science*, 15 (2): 135-175.

Kitcher, P., 1992, "The Naturalists Return", in *The Philosophical Review*, 101, 1, Philosophy in Review: Essays on Contemporary Philosophy: 53-114.

Kornblith, H., 1993, "Epistemic Normativity", in *Synthese*, 94 (3), 357-376.

Mitchell, S., 2009, "Complexity and Explanation in the Social Sciences", in C. Mantzavinos (ed.), *Philosophy of the Social Sciences: Philosophical Theory and Scientific Practice*, Cambridge University Press, Cambridge: 130-145.

Korsgaard, C., 1996, *The Sources of Normativity*, Cambridge University Press, Cambridge.

Korsgaard, C., 2009, *The Self-Constitution*, Oxford University Press, Oxford.

Leiter, B., 2004, "Introduction", in B. Leiter(ed), *The Future for Philosophy*, Oxford University Press, New York.

Macarthur, D. and DeCaro, M., 2004, "The Nature of Naturalism" in M. DeCaro and D. Macarthur (eds.), *Naturalism in Question*, Columbia University Press, New York: 1-20.

Macarthur, D., 2010, "Taking the Human Sciences Seriously", in M. DeCaro and D. Macarthur (eds.), *Naturalism and Normativity*, Columbia University Press, New York: 123-141.

Macarthur, D., DeCaro, M., 2010, "Introduction: Science, Naturalism and the Problem of Normativity" in M. DeCaro and D. Macarthur (eds.), *Naturalism and Normativity*, Columbia University Press, New York: 1-19.

Mackie, J.L., 1977, *Ethics: Inventing Right and Wrong*, Pelican Books.

McDowell, J., 1996, *Mind and World*, Harvard University Press, Cambridge MA.

McDowell, J., 1998a, "Two Sorts of Naturalism" in *Mind, Value and Reality*, Harvard University Press, Cambridge MA: 167-197.

McDowell, J., 1998b, "Functionalism and Anomalous Monism" in *Mind, Value and Reality*, Harvard University Press, Cambridge MA.: 325-340.

McDowell, J., 2000, "Responses", in *John McDowell: Reason and Nature*, edited by M. Willaschek, LIT-Verlag, Munster: 93-117.

McDowell, J., 2002, "Responses" in N.H. Smith (ed), *Reading McDowell on Mind and World*, Routledge, London-New York: 269-305.

McDowell, J., 2006, "Response to Graham Macdonald", in C. Macdonald & G. Macdonald (eds), *McDowell and his Critics*, Blackwell, Singapore: 235-239.

McDowell, J., 2008, "Responses"in Lindgaard, J. (ed) *John McDowell: Experience, Norm, and Nature*, Blackwell, Singapore: 200-267.

McDowell, J., 2009,"Naturalism in the Philosophy of Mind", in *The Engaged Intellect: Philosophical Essays*, Harvard University Press, Cambridge MA: 257-278.

McDowell, J., 2010, "Autonomy and its Burdens", in *The Harvard Review of Philosophy*, 17 (1): 4-15.

McDowell, J., 2018, "Responses", in F. Sanguinetti and A.J. Abath (eds.), *McDowell and Hegel: Perceptual Experience, Thought and Action*, Springer, Cham: 231-258.

Mitchell, S., 2009, "Complexity and Explanation in the Social Sciences", in Ch. Matzavinos (ed.), *Philosophy of Social Sciences: Philosophical Theory and Scientific Practice*, Cambridge University Press, Cambridge: 130-143.

Mizrahi, M., 2013, "The Pessimistic Induction: A Bad Argument Gone Too Far", in *Synthese*, 190 (15): 3209-3226.

Parfit, D., 2006, "Normativity", in R. Shafer-Landau (ed), *Oxford Studies in Metaethics Vol. 1*, Clarendon Press, Oxford: 325-380.

Pinkard, T., 2002, *German Philosophy 1760-1860: The Legacy of Idealism*, Cambridge University Press, Cambridge.

Pippin, R., 2008, *Hegel's Practical Philosophy*, Cambridge University Press, Cambridge.

Pippin, R., 2018, "A Response to McDowell on Hegel", in F. Sanguinetti and A.J. Abath (eds.), *McDowell and Hegel: Perceptual Experience, Thought and Action*, Springer, Cham: 211-230.

Rödl, S., 2007, *Self-Consciousness*, Harvard University Press, Cambridge MA.

Rorty, R., 2010, "Naturalism and Quietism", in M. DeCaro and D. Macarthur (eds.), *Naturalism in Question*, Columbia University Press, New York: 55-68.

Risjord, M., 1998, "Norms and Explanation in the Social Sciences", in *Studies in the History and Philosophy of Science*, 29 (2): 223-237.

Risjord, M., 2014, *Philosophy of Social Science: A Contemporary Introduction*, Routledge, New York-London.

Risjord, M., 2016, "Introduction", in M. Risjord (ed.), *Normativity and Naturalism in the Philosophy of the Social Sciences*, Routledge, New York-London: 1-8.

Rovane, C., 2010, "Why Scientific Realism May Invite Relativism", in M. DeCaro and D. Macarthur (eds.), *Naturalism in Question*, Columbia University Press, New York: 100-120.

Sellars, W., 1956/1997, *Empiricism and Philosophy of Mind*, Harvard University Press, Cambridge MA.

Rouse, J., 2016, "Toward a New Naturalism: Niche Construction, Conceptual Normativity, and Scientific Practice", in M. Risjord (ed.), *Normativity and Naturalism in the Philosophy of the Social Sciences*, Routledge, New York-London: 28-42.

Stueber, K.R., 2012, "Understanding Versus Explanation? How to Think about the Distinction between the Human and the Natural Sciences", in *Inquiry*, 55 (1): 17-32.

Turner, S., 2010, *Explaining the Normative*, Polity, Cambridge.

Turner, S., 2016, "The Naturalistic Moment in Normativism", in M. Risjord (ed.), *Normativity and Naturalism in the Philosophy of the Social Sciences*, Routledge, New York-London: 9-27.

Focus

Essays on Barbara Vetter's *Potentiality*

Introduction

Lorenzo Azzano, Andrea Borghini

1. *Introduction*

This focus of *Philosophical Inquiries* is devoted to Barbara Vetter's *Potentiality: From Dispositions to Modality* (Oxford University Press, 2015). In her book, Vetter offers an account of (certain) modalities in terms of the dispositions of certain entities; examples of these dispositions are a glass' fragility, or a rubber band's elasticity. More specifically, Vetter's account aims to explain metaphysical possibility and necessity in terms of a generalized notion of dispositionality. She refers to such notion as *potentiality*. We shall return on the nuances of Vetter's choice in section 3 below; for now, we shall speak more generally of 'dispositions' and 'dispositionality.'

Vetter is not the first to suggest a dispositional treatment of modality, not even in contemporary philosophy;[1] yet, *Potentiality* is surely the most articulated effort in that direction: as such, it has sparked the interest of many, friends and foes of dispositions alike. In addition, the relevance of *Potentiality* is not confined to modal ontology/semantics, or the somewhat narrow subfield of philosophy of science concerning dispositions, conditionals, laws of nature, and causation: on the contrary, the book presents significant ramifications in a vast array of debates, from metaphysics and modal epistemology to philosophy of mind, ethics, aesthetics, and more. Some of these ramifications are explored in the remainder of the present focus, which is intended as a contribution to an emerging debate about the complex interplay between dispositionality and modality.

The goal of our brief introductory remarks is, firstly, to frame the debate and answer some preliminary questions the reader might have (why a dispositional account of modality? and what is special about Vetter's own account?); secondly, we provide a brief overview of the history and contents of the focus itself.

[1] The first explicit discussion about a dispositional treatment of modality in contemporary analytic philosophy can be found in Mondadori&Morton 1976. More recently: Martin&Heil 1999; Pruss 2002; Molnar 2003: ch. 12; Mumford 2004: ch. 10; Borghini&Williams 2008; Contessa 2010; Jacobs 2010.

philinq VIII, 1-2020, pp. 71-82
ISSN (print) 2281-8618-ETS DOI: 10.4454/philinq.v8i1.275

2. *A background on dispositional treatments of modality*

The core idea of a dispositional account of modality is very simple. Consider a glass bottle. The bottle can break, and it would break if struck. How can such facts be explained? The answer of those who embrace a dispositional account is that one of such facts (or both, depending on the account) is grounded and explained by the fact that the glass bottle is disposed to break (if struck). More generally, so-called *dispositionalists* maintain that the modal features of the world entirely depend of instances of genuine dispositions by material objects in the physical world.

Some readers may find the dispositionalist project objectionable from the get-go. To say that the glass bottle is disposed to break (if struck) ultimately amounts to say that it *can* break, or that it would break if struck (or some elaboration thereof), thus making any attempt at a dispositional treatment of such modalities hopelessly circular. Then one should keep in mind that dispositional treatments of modality revert the traditional direction of analysis, according to which modal resources are deployed to explain dispositionality away, usually (but not exclusively) along the lines of a "conditional analysis": something is disposed to M, if S, *if and only if* it would M, if S.[2] On the contrary, friends of dispositions, capacities, and causal powers take them to have a life on their own, and to be primitive and irreducible items that the physical world and its inhabitants are provided with. Uncompromising claims such as this constitute a relative novelty in a philosophical landscape whose most celebrated influence was Hume's empiricism, according to which, "[o]f all the ideas that occur in metaphysics, none are more obscure and uncertain than those of power, force, energy or necessary connection".[3] Hume's opinion was later reinforced along verificationist lines: for how can the ascription of a dispositional predicate be meaningful, let alone true, when the presence of most dispositions, like the fragility of a glass bottle inside a cupboard, eludes standard verification procedures? Hence, the need for a conditional analysis.

The banishment and subsequent rehabilitation of dispositions is a complicated topic with multiple aspects to be considered. But it is indeed the position of many today that dispositions are respectable items that do not need crutches of any kind (be they semantic or metaphysical) to stand upright;[4] and

[2] Traditionally, Carnap 1936-37; Ryle 1949; Goodman 1954; Quine 1960. Unlike most recent attempts, as in Lewis 1997, original supporters of a conditional analysis, starting with Carnap, limited themselves to extensional logic.

[3] From *An Enquiry Concerning Human Understanding*, Section 7, Part I.

[4] Perhaps the first explicit contemporary manifesto in favor of dispositions can be found in Mellor 1974; Martin 1994 and Bird 1998 counterexamples to the conditional analysis also indubitably

it is this position which makes a dispositional treatment of modality viable.

Of course, however, that a philosophical option is viable doesn't mean that it is desirable: why, then, pursue a dispositional treatment of modality? The concerns put forward by dispositionalists in this context are most often than not ontological in character: dispositions are better candidates for an ontology of modality than the far more popular *possible worlds*.

The topic warrants some elaboration. As soon as modal ontology became a respectable topic, the debate has been monopolized by discussions about possible worlds.[5] The reader should be warned that interest in possible worlds in modal ontology has little to do with Leibniz's theological considerations, or with the so-called many-world interpretation of quantum mechanics: it rather has to do with the theoretical virtues of the modal semantics it is based on. Possible-world modal semantics is, as of now, the best way to understand and extensionalize modal discourse: non-truth-functional modal sentences are understood through a truth-functional semantic metalanguage, and otherwise opaque modal operators are interpreted as well-understood quantifiers; what is more, many features of modal logic can be nicely framed through possible world models (famously, by tinkering the mathematical features of the accessibility relations, modal logics of various strength are validated). So, the thought goes, if modal discourse is at least sometimes true, and true in the same way in which non-modal discourse is at least sometimes true, then the best semantics that can be provided for it ought to correctly describe reality; hence, possible worlds.[6] The alternative is to take a fictionalist, or non-cognitivist, or otherwise non-realist stance on the status of modal discourse.

Dispositionalists are aware of this difficulty; Jacobs (2010: 240) is quick to recognize that a pressing matter for dispositionalists is to develop an "alternative to the powerful, possible worlds semantics of modality". There are, however, other factors at play. Firstly, one should consider that this allegedly virtuous possible-world formal semantics does not straightforwardly involve possible worlds, at least not without an active interpretative effort.[7] Kripkean triples <W,@,R> are merely characterized by set-theoretical features, and it is then philosophers who decide to interpret them through the lenses of a possible world ontology. Possible world talk is surely a useful and comfortable way

played a part.

[5] The most radical and popular version of the position is of course Lewis 1986; see Divers 2002 for a comprehensive overview on the subject.

[6] See Mondadori&Morton (1976).

[7] The distinction between pure and applied semantics is relevant here, as in Plantinga (1974: 126ff). Relatedly, Fleischer (1984) described Kripkean modal semantics (somewhat provocatively) as an amalgamation of algebra and poetry. Possible worlds classify as poetry.

to frame modal discourse (and one which Vetter herself employs in the book), but one whose ontological commitments are *per se* dubious.

Secondly, it is not clear whether the discussion in modal ontology should be completely stifled by semantic considerations; on the contrary, it could be problematic to read off one's ontology from the formal semantics alone. If an ontology of dispositions can be shown to be preferable to possible worlds for inherently ontological reasons, then a discussion about the semantics may follow it, rather than precede it. This is the kind of methodology put forward in Borghini&Williams (2008), who however do not go as far as so to propose a semantic for their disposition-based modal ontology. Vetter's *Potentiality* does both, ontology and semantics, although the comparison between possible world semantics and her potentiality semantics is deferred to another time.

What are, then, the ontological reasons to prefer dispositions in modal ontology? There's no clear-cut list of uncontroversial virtues enjoyed by dispositions in opposition to possible worlds; there is however, a number of more or less interconnected observations put forward by dispositionalists throughout the literature.

Firstly, actualism is often stated to be a virtue of dispositional treatments of modality, to the extent that they ground modal features in actual instances of dispositions.[8] That said, there are actualist possible world ontologies as well, so this cannot be the decisive factor to make us decide in favor of dispositions over possible worlds.[9] Secondly, it is sometimes claimed that possible worlds, whether actual or not, are wholly irrelevant when producing an account of modality: that the glass bottle is represented as broken in an *ersatz* world, or has a broken counterpart in another genuine world, doesn't seem to be the reason why the glass bottle *can* break in the first place; of course, the charge of irrelevance against possible worlds (and counterpart theory) is not a new one,[10] but one which many dispositionalists, Vetter included, take seriously.[11] The background idea is that the only reason why the glass bottle can break is because of some perfectly respectable property that it possesses, which can be picked by the predicate "fragile".

This leads us to our third point concerning dispositions *qua* modal ontology: that they allow for a thoroughly *naturalistic* ontology, at least in the Armstrongian sense of naturalism as "the doctrine that reality consists of nothing

[8] Eagle 2009; Contessa 2010; and Vetter 2011.

[9] Contessa (2010) differentiates traditional actualism (e.g., the one present in actualist possible world ontologies) from "hardcore actualism" which rejects the idea that the so-called "Leibnizian biconditionals" employed in possible world semantics ought to be taken ontologically seriously.

[10] For an overview on this topic, see Divers (2002: 124-133).

[11] See Jacobs 2010 and *Potentiality*: 6.

but a single all-embracing spatio-temporal system".[12] To ground modal features of the world one only needs perfectly respectable instances of properties within the space-time continuum, without the need to explain them away through far-removed possible worlds which are only postulated by virtue of the corresponding formal semantics.

Fourthly, and finally, there's an epistemological advantage. Peacocke (1999: 1) called "integration challenge" the general task of providing "for a given area, a simultaneously acceptable metaphysics and epistemology". This has been a notoriously problematic aspect of possible world accounts of modality, the most glaring issue being that there is no epistemic access to possible worlds, direct or indirect as it might be, since, as Kripke's adagio goes, we can't see them with a telescope. But there *is* a fairly straightforward access to manifestations of dispositions (for most of them, we don't need telescopes either!); which means that a dispositional treatment of modality offers the prospect of a fruitful integration between modal ontology and modal epistemology.

3. *Vetter's approach in* Potentiality

As before, Vetter's variant of dispositionalism is not the only one on the market. What is specific, then, to her approach in *Potentiality*, and which motivations are offered in its favor?[13]

Firstly, we must understand Vetter's distinctive shift from dispositions to *potentialities*. Dispositions, such as a sugarcube's water solubility, are not the only features associated with what entities can or would do in certain circumstances. Entities might also possess certain powers, capacities, or potentials, e.g. water's power to dissolve sugar;[14] in addition, some entities have abilities, such as Jane Austen's ability to write in English. Vetter intends "potentialities" (a term of art which she hopes to be conceptually blank) to be the most general category in which they all fit; potentiality is thus intended to be the "common genus" (102) of dispositions and other modal features.

The one just described is not the only respect in which dispositions and potentialities differ. Dispositional ascriptions are notably a messy affair, subject to both vagueness and context-sensitivity: e.g., a XVII century Ming vase

[12] Armstrong 1981: 149.

[13] Numbers in brackets, unless otherwise specified, refer to *Potentiality*.

[14] The distinction in many natural languages between "dispositions" and "powers" may be linked to the pre-theoretical asymmetry (of ultimately Aristotelian descent) between an agent and a patient in a causal process, such as the dissolution of a sugarcube in water. Many friends of dispositions today believe this asymmetry to cut no ontological ice.

probably counts as fragile, and has to be bubble-wrapped wherever it goes, whereas a steel rod appears to be pretty sturdy, thus not fragile; but what about a pencil? I can break it with one hand, but I do not need to be excessively careful handling it in my everyday life, since it can probably survive a fall from my desk without breaking: is that enough to make it fragile? As for context-sensitivity, it is not even uncontroversial that a steel rod is *not* fragile: a certain steel rod may be considered to be too fragile by engineers on a construction site, and thus not used for a skyscraper's foundation. However, according to Vetter, "both vagueness and context-sensitivity are features of language, not the world" (20); thus potentialities are introduced as a context-neutral metaphysical background for dispositional ascriptions. Potentialities come in degrees: both the Ming vase and the steel rod have a potentiality to break, but with a different degree, thus, warranting, in different contexts, different dispositional ascriptions. In many ordinary contexts the degree of the Ming vase's potentiality passes a threshold for the vase being called "fragile", while this is not the case for the steel rod; but in other contexts, different dispositional ascriptions might be warranted.

Here's a different way to make sense of the vagueness and context-sensitivity of dispositional ascriptions: viz., to specify a different stimulating condition. After all, the Ming vase is more fragile than the steel rod to the extent that it takes a far lesser force to shatter it. Thus, the Ming vase has a disposition to break when struck with (at least) *n*, whereas the steel rod has a disposition to break when struck with (at least) *m*, such that n<m. This brings us closer to the so-called Standard Account of dispositions, in which dispositions are characterized and individuated by virtue of a stimulus S and manifestation M, and usually paired with the correspondent conditional "if S then M".

Vetter takes this option to be deeply problematic. She argues (39-49) that no disposition, unless maximally specific ones (e.g., disposed to break if struck with *exactly* 8.5 N), can successfully be paired to a single stimulus and manifestation, thus to a single conditional (so-called "single-track" dispositions, as opposed to the "multi-track" ones). Thus, a general disposition such as fragility would have to be understood as a massively multi-track, or as complex disjunction of maximally specific dispositions; intuitively, that general disposition would then be less fundamental than the maximally specific dispositions (in the same way a complex disjunction is less fundamental than its disjuncts);[15] however, Vetter also has an argument to the contrary conclusion that general dispositions are more fundamental than maximally specific ones (56-58).

[15] See Bird 2007: 22.

The culprit, for Vetter, is the Standard Account itself, which characterizes the nature of a disposition in terms of both stimulus and manifestation; her idea (63 ff) is to characterize it by its manifestation alone, and take the resultant potentiality to come in degrees, rather than the stimulus.[16] This solves the difficulty: the Ming vase and the steel rod both possess the general potentiality to break, but to different degrees. Potentialities, so understood, are a modal primitive of the account, and are thus to be taken as fundamental (or, at least, no less fundamental than anything else, 24-25).

This kind of internal restructuring of dispositionality has a crucial consequence on Vetter's account of modality; if the Standard Account naturally pairs dispositions with conditionals (from "object *a* is disposed to M if S" to "object *a* would M if S"), Vetter's potentialities are naturally paired with possibility (from "object *a* has a potentiality to break" to "object *a* can break"); thus, *Potentiality* offers a *possibility-first*, rather than *conditional-first* account of modality.[17] An immediate advantage is that Vetter is spared from having to deal with the notoriously troublesome link between dispositional ascriptions and conditionals.

In order to produce a formal semantics, Vetter introduces her potentiality operator POT (144-145), which functions as a predicate modifier; this is in line with the natural language grammar of dispositionality (e.g., disposed to break), but in order to bridge the difference in logical form between POT and possibility, *qua* sentential operator, Vetter allows predicate abstraction to turn sentences of any given logical complexity into predicates to plug in the POT operator; this is, by Vetter's own admission "the path of least formal resistance" (141). Vetter introduces multiple kinds of potentiality to ensure that all such applications of the POT operator make sense from an ontological point of view; from the simplest intrinsic potentialities to *joint potentialities*, viz. potentialities jointly possessed by two or more items (e.g., Lorenzo's and Andrea's joint potentiality to play a game of chess), to *extrinsic potentialities* (e.g., Lorenzo's potentiality to play a game of chess with Andrea, or maybe even Lorenzo's potentiality for Andrea to play a game of chess). Finally, building from Borghini&Williams (2008) there are *iterated potentialities*, viz., potentialities whose manifestations are, or involve, the possession of other potentialities (e.g., Lorenzo's potentiality to learn to play chess). Vetter claims extrinsic potentialities to be grounded in joint potentialities: as long

[16] Reasons for her alternative account also come from linguistic considerations concerning dispositionality in natural languages; for according to Vetter "our modal metaphysics should provide the materials for a semantics of at least a significant part of natural-language modality" (16).

[17] See Jacobs 2010 for a conditional-first account of modality. Even without potentialities in Vetter's sense, Borghini&Williams 2008 also argue for a possibility-first account.

as one can produce a reasonable grounding chain ending in the simplest intrinsic potentialities, every potentiality is acceptable.

This ample variety of potentialities greatly expands the reach of potentiality, and gives ontological significance to all applications of the POT operator. The criterion that POT must respect in this endeavor is some degree of "extensional correctness" (15ff.), it must respect enough of our pre-theoretical intuitions about what is possible, and what is not possible.

Finally, in page 197, Vetter presents the POSSIBILITY principle:

POSSIBILITY It is possible that *p* =df Something has an iterated potentiality for it to be the case that *p*.

Necessity is standardly introduced as the dual of possibility: thus, it is necessary that *p* if and only if nothing has the potentiality to not-*p* (203). As Vetter quickly points out, POSSIBILITY as a proper definition could be used to replace and eliminate all possibility-talk, but it shouldn't. The account doesn't offer a reduction of modality in the sense that modal features of the world are reconducted to non-modal features: it rather consists in a *localization* of modality to the specific aspects of the physical world which are responsible for it (viz., instances of potentialities). In POSSIBILITY such a localization is signaled by the existential quantification in the right-hand side. On the other hand, as Vetter claims (197-198), possibility-talk involves some kind of abstraction, as we are required to think about some potentiality in abstraction from its bearer.

4. *About this focus*

The idea of this focus presented itself quite naturally following two events centered on *Potentiality*. A reading group on Vetter's book met for an extensive seminar at the University of Padua, on July 7, 2017. The meeting was organized by Massimiliano Carrara, Giorgio Lando, and Vittorio Morato and saw the participation of Lorenzo Azzano, Massimiliano Carrara, Donatella Donati, Ciro De Florio, Simone Gozzano, Giorgio Lando, Vittorio Morato, Alessio Santelli, Alfredo Tomasetta, and Giacomo Turbanti. The reading group was later followed by the workshop *Potentiality & Possibility*, organized by Giorgio Lando at the University of L' Aquila on September 14, 2017, with talks from Barbara Vetter, Andrea Borghini, Lorenzo Azzano, and Donatella Donati.

The focus hosts six contributions from seven philosophers, engaging with different aspects of Vetter's book, in an attempt to advance the discussion on potentialities, and, more generally, on dispositional treatments of modality. In the last part, Vetter offers her reply to the contributors.

The first paper is *Possibility and the Analysis of Dispositions* by Alexander Bird, which studies the possibility-first aspect of Vetter's treatment of modality. Linking potentialities directly with possibility *prima facie* spares Vetter from a problem of conditional-first approaches: that of non-conditional possibilities. According to this problem (which is fairly new to the literature), when dispositions are provided with both stimulus and manifestation, and thus paired with subjunctive conditionals, the non-conditional possibility of the consequent can only be ensured through the non-conditional possibility of the antecedent, which no disposition can provide. It would seem that Vetter has the upper hand here; unfortunately, she allows for *conditional manifestations* in order to deal with those dispositions that would normally be distinguished by their stimulus (e.g., gravitational mass and electric charge). Although technically stimulus-less, potentialities with conditional manifestations can be shown to be subject to the problem of non-conditional possibilities.

David Yates, in *A Strange Kind of Power*, questions the formal adequacy of Vetter's account. Yates had previously argued (Yates 2015) that, given Vetter's account, for certain propositions *p* (e.g., 2+2=4), it is both the case that *p* is necessary and not possible; given minimal requirements for the formal adequacy of the account, (axiom T) that yields a contradiction.[18] Vetter has later argued for a plenitude of potentiality, according to which there can be a non-causal potentiality for it to be the case that, e.g., 2+2=4 (Vetter 2018). Yates criticizes Vetter's strategy for plenitude, which crucially revolves around the claim that if an object necessarily has an intrinsic property P, then it also is maximally disposed to P. An alternative is offered: by treating truthmaking as a form of metaphysical causation, that something makes 2+2=4 true suggests that it may also possess a causal potentiality to do so: pros and cons for that "strange kind of powers" are then evaluated.

The two following papers deal with an important issue in the metaphysics of potentialities: assuming that potentialities are indeed properties, as Vetter does throughout her book, what kind of properties are they? Between various brands of nominalism, trope theory, and universalism, which account of properties, if any, best fits *Potentiality*? The two papers both find Vetter's current stance on the matter somewhat unstable, but advocate for a different solution: Platonist the first, Aristotelian the second.

In *Potentiality: Actualism minus Naturalism equals Platonism*, Giacomo Giannini and Matthew Tugby suggest that Vetter's potentiality account would be better served by a Platonist framework. Vetter takes the ontology of potentialities to be appealing insofar as it is both actualist and naturalist (as defined

[18] For these formal requirements, and T in particular, in *Potentiality*, see 15-16.

above); this leads her to an Aristotelian account of properties as immanent universals, which depend on their instances and thus cannot exist uninstantiated. Yet this dependence notoriously narrows the range of possibilities that can be accounted for. Vetter is aware of this difficulty[19], and in *Potentiality* she formulates a weakening of her position (271-272); but for Giannini and Tugby, in this weakening the naturalist component of Vetter's ontology is already compromised. It is thus a small step to abandon the dependence claim and embrace Platonism, according to which properties can exist uninstantiated. The advantage is that there are certain aspects of scientific reasoning that are better understood through a Platonic framework.

Potentialities as Properties by Jennifer McKitrick has a wider scope, arguing that no existent theory of properties is compatible with Vetter's account: all accounts are problematic, including Platonism. So rather than following Giannini and Tugby in an explicit rejection of naturalism, McKitrick focuses on a different aspect of Vetter's framework to solve this difficulty: viz., the claim that determinable potentialities are more fundamental than their determinates, to which Vetter may be committed, according to McKitrick, by virtue of her claim that general potentialities are more fundamental than the specific ones. However, according to McKitrick, *Potentiality* may also offer the solution: building from Vetter's tentative suggestion that so-called nomological dispositions, which "encode laws of nature" (*Potentiality*, 50), e.g. electric charge, are always possessed to the maximal degree, she claims fundamental potentialities to be such determinate nomological potentialities. This alternative proposal is consistent with many options about the status of properties (Aristotelianism, and perhaps trope theory as well).

Nathan Wildman also offers three *Potential Problems?* relating to potentialities. Firstly, regarding talk of degrees of potentiality, for Vetter to be cashed out with the help of a proportionality principle formulated in terms of possible worlds. Such a principle should of course not be taken ontologically too seriously, but merely as a "formal model and rough approximation' of degrees of potentiality" (*Potentiality*, 78); thus, at least, materially adequate. Wildman offers a counterexample; Vetter can of course abandon the problematic principle entirely, but that would leave talk of degrees as entirely primitive. Secondly, there's a difficulty concerning the individuation of potentialities; given that Vetter rejects stimuli, all the individuating work has to be done by manifestations alone; but Wildman objects that there might be different potentialities with the same manifestations, which apply to different kinds of entities (e.g.,

[19] Limitations of an Aristotelian (and trope theorist) account of dispositions are forcefully pressed in Tugby 2013.

perishable and destructible). Thirdly, Wildman wonders how Vetter could treat potentialities whose manifestation involve the bearer's ceasing to exist (e.g., again, perishability and destructibility).

In the last contributed paper *Potentiality, Modality, and Time*, Jennifer Wang introduces a distinction between "*de re* first" treatments of modality, which start from modal properties of objects, and "*de dicto* first" treatments, which start from general possibilities and necessities. Vetter's account, as exemplified by the POT predicate modifier, is of the former kind, while possible world ontologies (at least in the paradigmatic Lewisian variant) clearly belong to the latter. Vetter's intended treatment for *de dicto* modalizations, e.g., possibly the US president is a woman, requires backtracking in time in search for bearers of potentialities whose manifestation might have brought about that, say, the US president is a woman. Wang objects that, in the case of some specific *de dicto* modal claims, Vetter's account loses its attractiveness and intuitiveness. Yet *de dicto* first treatments do not necessarily need possible worlds to function: an alternative *de dicto* first account is also discussed, with primitive modal relations of compatibility and incompatibility between properting; eventually, Wang suggests that a mixed account might serve Vetter better.

Lorenzo Azzano
lorenzo.azzano@unibg.it
Università di Bergamo

Andrea Borghini
andrea.borghini@unimi.it
Università degli Studi di Milano

References

Armstrong, D.M., 1981, *The Nature of Mind and Other Essays*, Cornell University Press, Ithaca.

Bird, A., 1998, "Dispositions and antidotes", in *Philosophical Quarterly*, 48 (191): 227-234.

Bird, A., 2007, *Nature's Metaphysics: Laws and Properties*, Oxford University Press, Oxford.

Borghini, A., & Williams, N.E., 2008, "A dispositional theory of possibility", in *Dialectica*, 62 (1): 21-41.

Carnap, R., 1936-1937, "Testability and meaning", in *Philosophy of Science*, 3 (4): 419-471.

Contessa, G., 2010, "Modal truthmakers and two varieties of actualism", *Synthese*, 174 (3): 341-353.

Divers, J., 2002, *Possible Worlds*, Routledge, London.

Eagle, A., 2009, "Causal structuralism, dispositional actualism, and counterfactual conditionals", in T. Handfield (ed.), *Dispositions and causes*, Oxford University Press, Oxford: 65-99.

Goodman, N., 1954, *Facts, Fiction and Forecast*, University of London Press, London.

Hume, D., *An Enquiry Concerning Human Understanding.*

Fleisher, I., 1984, "Kripke Semantics = Algebra + Poetry", in *Logique et Analyse*, 27: 283-295.

Jacobs, J., 2010, "A powers theory of modality: or, how I learned to stop worrying and reject possible worlds", in *Philosophical Studies*, 151 (2): 227-248.

Lewis, D.K., 1986, *On the Plurality of Worlds*, Wiley-Blackwell, Hoboken NJ.

Lewis, D.K., 1997, "Finkish dispositions", in *Philosophical Quarterly*, 47 (187): 143-158.

Martin, C.B., 1994, "Dispositions and conditionals", in *Philosophical Quarterly*, 44 (174): 1-8.

Martin, C.B., & Heil, J., 1999, "The ontological turn", in *Midwest Studies in Philosophy*, 23 (1): 34-60.

Mellor, D.H., 1974, "In defense of dispositions", in *Philosophical Review*, 83 (2): 157-181.

Molnar, G., 2003, *Powers: A Study in Metaphysics*, Oxford University Press, Oxford.

Mondadori, F., & Morton, A., 1976, "Modal realism: The poisoned pawn", in *Philosophical Review*, 85 (1): 3-20.

Mumford, S., 2004, *Laws in Nature*, Routledge, London.

Peacocke, C., 1999, *Being Known*, Oxford University Press, Oxford.

Plantinga, A., 1974, *The Nature of Necessity*, Clarendon Press, Oxford.

Pruss, A.R., 2002, "The actual and the possible", in R.M. Gale (ed.), *The Blackwell Guide to Metaphysics*, Blackwell, Oxford: 317-333.

Ryle, G., 1949, *The Concept of Mind*, Hutchinson and Co, London.

Quine, W.V., 1960, *Word and Object*, MIT Press.

Tugby, M., 2013, "Platonic dispositionalism", in *Mind*, 122 (486): 451-480.

Vetter, B., 2011, "Recent Work: Modality without Possible Worlds", in *Analysis*, 71 (4): 742-754.

Vetter, B., 2015, *Potentiality: From Dispositions to Modality*, Oxford University Press, Oxford.

Vetter, B., 2018, "A plenitude of powers", in *Synthese*, DOI 10.1007/s11229-018-1758-z

Yates, D., 2015, Dispositionalism and the Modal Operators, *Philosophy and Phenomenological Research*, 91 (2): 411-424.

Possibility and the analysis of dispositions

Alexander Bird

Abstract: I examine Barbara Vetter's dispositional account of modality and the analysis of dispositions upon which it is based. The latter ties dispositions to manifestations only. I argue that this feature gives Vetter's account an advantage over other dispositional accounts of modality – it avoids the 'problem of [how to ground] non-conditional possibilities'. On the other hand, I argue that we need stimuli as well as manifestations in order to distinguish distinct dispositions that have identical manifestations. Vetter's answer to the latter says that some manifestations are conditional in nature. That answer undermines the advantage that Vetter's account had with regard to the problem of non-conditional possibilities – her view now also faces that problem. I also raise the question of how it is that the analysis of an 'everyday' concept such as 'disposition' could provide insight into fundamental questions of modal metaphysics.

Keywords: disposition; modality; possibility; the problem of non-conditional possibilities; Vetter.

1. *Introduction – a new account of dispositions*

Dispositions have modal characteristics. They have implications for what could or would happen. If dispositions – or, better, properties with dispositional natures (dispositional properties, for short) – are a fundamental part of our ontology, might not their existence provide the basis for an account of modality? While I have (Bird 2007: 218 fn143) made such a suggestion, others have actually attempted the task (Borghini and Williams 2008; Jacobs 2010). Barbara Vetter's (2015) wonderfully rich book gives the most thoroughly detailed dispositional account of modality yet available. It does much else besides, but crucially it gives a novel account of what dispositions themselves are, and this forms the basis for her account of modality.

The standard account of dispositions (SA) says that something like the following is true:

philinq VIII, 1-2020, pp. 83-96
ISSN (print) 2281-8618-ETS DOI: 10.4454/philinq.v8i1.276

(SA) for a disposition D, and appropriate stimulus condition S and manifestation condition M, *x* has D iff were *x* to be S then *x* would M.

For example:

(F-S) *x* is fragile iff were *x* subjected to a stress, *x* would break.

(SA) may be strictly false because of finks, antidotes/masks and the like (Johnston 1992; Martin 1994; Lewis 1997; Bird 1998). These are conditions that interfere with the normal course of events, so that the truth of the subjunctive 'were *x* subjected to a stress, *x* would break' does not align with the truth of the disposition ascription '*x* is fragile'. A stimulus to a disposition might in special circumstances cause the disposition to disappear. For example, striking a fragile glass might also cause it to be superheated very quickly so that the glass becomes soft and pliable before the striking can lead to its breaking. So, although it is fragile, striking the glass would not cause it to break. In the case of a mask or antidote the stimulus does not eliminate the disposition but does initiate interference with the normal course of the disposition's action. A bite from a deadly coral snake is disposed to kill me. But I take an antidote and thereby survive. These show that in some circumstances there can be dispositions without corresponding true subjunctive conditionals. Likewise, there can be true subjunctive conditionals without dispositions. At the moment the glass is soft. But were I to strike it, it would be rapidly supercooled and so become fragile and break from that very striking. A sturdy iron cooking pot is attached to a bomb with a sensitive detonator. The pot is not fragile. But were I to strike it, it would end up broken. The latter is an example of a *mimic* to a disposition.

Even though such cases show (SA) to be strictly false, most commentators think that there is something fundamentally correct about (SA). After all, the cases mentioned in the preceding paragraph are unusual. Some hold that (SA) can be patched up, with clauses added to exclude these interferers. Others holds that even if (SA) cannot be made watertight, it is nonetheless close to the truth. In any case, different versions of the standard view of dispositions all agree that the characterisation of a disposition requires specification of both a stimulus condition and a manifestation (perhaps more than one of each).

In contrast with the standard view, Vetter holds that dispositions should characterized solely in terms of their manifestation conditions. Hence Vetter's account (VA) says:

(VA) for a disposition D, and appropriate manifestation condition M, *x* has D iff *x* could M.

For example:

(F-V) *x* if fragile iff *x* could break.

(In fact Vetter says: *x* if fragile iff *x* could break *easily*. The latter addition, required to address the fact that sturdy things could break is stressed enough, is important and will be addressed later.)

This essay addresses two questions. First, is the new view of dispositions is correct? And, secondly, what implications does our answer have for Vetter's account of modality? The argument of the paper is that, for all her insights, there are problems with Vetter's accounts of dispositions and of modality. I first introduce the 'problem of non-conditional possibilities'. This I argue is a problem for standard approaches that account for modality in terms of dispositions. Vetter's approach seems not to suffer from this problem, and so her account has a prima facie advantage. I then look at her account of dispositions in terms of manifestations alone, on which the account of modality depends. In the sections 'The need for stimuli' and 'The problems of distinct dispositions' I give two arguments for thinking that a satisfactory account of dispositions needs stimuli, not manifestations alone. There is in Vetter's armoury a response to the second, more serious problem – she allows for conditional manifestations. I argue, however, that committing to this response means that Vetter's account will also suffer from the problem of non-conditional possibilities. So not only does her account lose its advantage over the standard view, it thereby also faces what I take to be a major obstacle to a dispositional account of modality.

2. *The problem of non-conditional possibilities*

Vetter's account (VA) makes the connection between dispositional properties and modality at least superficially straightforward (though there is a lot of detail that Vetter also provides). Borghini and Williams (2008) also propose to account for modality in terms of dispositions. But they use the standard account, (SA). That leads to what I regard as a significant problem with their approach (the problem of non-conditional possibilities). So, in my mind, one great advantage of (VA) is that it avoids this problem that is generated by using (SA).

As I say, we can see quite easily how to get the bare bones of an account of modality from Vetter's (VA):

(P) M is possible if something has a disposition D with manifestation M.[1]

[1] Vetter's account of possibility is much more sophisticated than this. See her chapters 5 and 6.

How would we get an account of modality from the traditional (SA)? (P) is not avilable. For the truth of (SA) does not make (P) true. The conditional $X\Box\!\!\rightarrow Y$ tells us that in the nearest possible world where X is true, Y is true also. But that makes Y true at some possible world only if X is true at some possible world. So, (SA) gives us only:

(P*) M is possible if something has a disposition D with manifestation M and stimulus S, and S is possible.

As a reductive account of possibility, (P*) does not look very promising on account of its circularity: (P*) tells us what it is for M to be possible in terms of S being possible (plus the dispositional claim). Similarly, S being possible will be articulated by a distinct instance of (P*), which will refer to a disposition D' of which S is the manifestation, and of which some other condition, T, is the stimulus – plus T being possible. And so on. This approach to understanding possibility in terms of dispositions will not bottom out in dispositions alone, but will always have an ungrounded appeal to some possibility.

Let us distinguish between non-conditional possibilities ('possibly, the glass is struck', 'possibly, the glass breaks') and conditional possibilities ('possibly, if the glass is struck, it breaks'). With (SA) we can ground conditional possibilities in dispositions. Our problem is that we cannot ground *all* non-conditional possibilities in dispositions. We can ground some non-conditional possibilities in dispositions, but only if there are some non-conditional possibilities that are not so grounded.

We can put this more formally. Let us assume (SA) and that *x* has disposition D. Then:

$$Sx\Box\!\!\rightarrow Mx$$

from which follows:

$$\Diamond(Sx \rightarrow Mx)$$

So we get a conditional possibility, such as its being possible that if *x* is struck *x* will break. But we also want non-conditional possibilities, such as $\Diamond Mx$. That is, in addition to its being possible that, if the vase it struck, it breaks, we also want it to be possible that the vase is broken. But we cannot derive such non-conditional possibilities from conditional ones of the form $\Diamond(Sx \rightarrow Mx)$. To derive $\Diamond Mx$ we also need $\Diamond Sx$ as a premise (in fact we need the compossibility of Sx and $Sx \rightarrow Mx$, i.e. $\Diamond(Sx \,\&\, Sx \rightarrow Mx)$). That is, to get to the (non-conditional) possibility that the vase is broken, we need not only the (conditional) possibility that if the vase is struck, it is broken, we need also the

(non-conditional) possibility that the vase is struck. But that is just to say that we can get an account of one non-conditional possibility only if we *already* have some other non-conditional possibility. I.e. this attempt to account for possibility is doomed to circularity (or infinite regress).

This problem arises because on the standard view the natures of all dispositional properties are conditional, and so they will only ever generate truths of the form of (P*) that are conditional in nature. Whereas a materially adequate account requires that we have truthmakers for non-conditional possibilities. Let's call this *the problem of non-conditional possibilities*. Giving a reductive account of modality in terms of dispositions while employing the standard account of dispositions suffers from the problem of non-conditional possibilities.

On the other hand, prima facie, Vetter's account of modality does not suffer from the problem of non-conditional possibilities. That is because Vetter's (VA) says that the natures of dispositions are non-conditional (or so it seems). For example, Vetter's account of fragility, (F-V), tells us that something is fragile if and only if it could break – we get the possibility of breaking directly from the disposition. More generally and formally, from (VA) and the assumption that *x* has disposition D, we derive directly that *x* could M and hence $\Diamond$M*x*.

So Vetter's approach to dispositions has, in my view, a considerable advantage over the standard view when it comes to providing a foundation for modality. Nevertheless, I shall argue that Vetter's view *also* turns out to suffer from a version of the problem of non-conditional possibilities.

3. *The need for stimuli*

Vetter's account has a significant prima facie virtue in accounting for possibility. But is that account right? In this section and in the next I articulate reasons for thinking that we cannot do without stimulus conditions in our account of dispositions.

Let's look at (F-V):

(F-V) *x* if fragile iff *x* could break.

That's clearly not quite right. For most things could break if stressed enough (e.g. subject to a huge explosive power), including many sturdy, not fragile things. So Vetter prefers something like:

(F-V') *x* if fragile iff *x* could break easily.

How is the 'easily' qualification supposed to be understood? It is to be understood, as others have done, in terms of close possible worlds.[2] Roughly, E could easily happen if it happens in some close possible world(s). The glass, being near the edge of the table, could easily have been knocked onto the floor: there is a world close to the actual world, differing only slightly in some earlier condition, in which the glass is in fact knocked onto the floor. For Vetter's purposes 'easily' implies more than just one possible world, to avoid the problem of Manley and Wasserman's (2008: 67) sturdy concrete block that would break if dropped in one precise way, but not otherwise. On the other hand, there do not need to be *many* worlds. As Vetter (2015: 73) says, the fragile and precious champagne glass might be carefully packed away at the back of a shelf. While the precarious glass on the table's edge is knocked off and broken in many worlds similar to the actual world, there are not many worlds where the precious champagne glass is unpacked and suffers breakage. Even so, there are 'a few' such worlds.

This understanding of 'easily' will not do, however. Let us exaggerate the champagne glass story. The wealthy lover of fragile glasses pays for the creation of a fragile glass under very controlled circumstances, and then ensures that it is packed in polystyrene (styrofoam), in a vault in an area at no risk of earthquakes, but with shock-proof engineering, protected by dedicated guards, and so forth. If we elaborate the story enough, it is clear that there is no close possible world in which this glass breaks (it comes to the end of its existence by melting). Nonetheless, it is intrinsically like many other fragile glasses – it is undeniably fragile. If one prefers a more realistic story, think of a high-security lab working with a dangerous infectious micro-organism. The organism might be infectious (a disposition), but in no close world does it infect anyone, thanks to the stringent precautions taken in the actual world. Indeed, the very point of bio-security is to make it the case that the dangerous organism cannot easily infect anyone. Yes, it is still highly infectious.

I think that Vetter has wrongly transferred an analysis of 'easily' appropriate for single events to the generic case of the manifestation of dispositions. There is a difference between these two cases:

(a) A glass is on the edge of the table; there are lots of people walking by; it could easily be knocked off, which would cause it to break. *The glass could easily be broken.*

[2] What determines the closeness of possible words? Vetter, who at this point in the discussion, is drawing in large part on the work of Angelika Kratzer (1981), is using a standard Lewisian conception of closeness, which in this case (where laws are held fixed) means that two worlds are close to the extent that they have exact matches in matters of particular fact.

(b) A glass is fragile; very little force would be required to snap its slender stem. *The glass could be broken easily.*

Only the first concerns happenings in close possible worlds. Regarding (a), the way to make the relevant (italicized) statement false is to move the glass away from the edge of the table, so that there is no longer a close world in which it is knocked off the table. In the second case, making the italicized statement false is not so easy, and no amount of careful positioning or other protection will help – one needs to change the glass itself. The two sentences are not interchangeable. One cannot use 'the glass could be broken easily' to mean that that it is actually in danger of being broken – one has to use the different 'the glass could easily be broken'.[3]

In (b) the role of 'easily' is to capture the thought that only a little stress needs to be applied to the glass for it to break. It does not tell us whether there is any prospect of the glass actually being stressed by that amount and therefore broken (there might not be, as in the case of the glass belonging to the lover of fragile things). That is, 'easily' relates to the implicit stimulus, a stressing of the glass. In short, in (b) to understand 'easily' we need to talk about the degree of stress that will bring about breaking, and to talk about stress is to talk about the stimulus. Consequently, I do not think that Vetter's (F-V) or (F-V'), which exclude any reference to a stimulus, can be right. Dispositions such as fragility must be characterised by a stimulus as well as a manifestation. That is true even if usually our interest is in the manifestation and so the stimulus will be left implicit or unspecified (we shall see that this is not always the case).

I have just argued that because 'easily' in (b) characterises the stimulus that brings about breaking, Vetter's stimulus-free account of dispositions cannot be right. I suggest that she has wrongly assimilated this 'easily' to the 'easily' of (a) which does not concern any stimulus but concerns nearby possible worlds. That mistaken assimilation means that (F-V') can be shown to be false by cases such as that of the sturdy iron pot attached to a bomb with a sensitive detonator. Just like the precarious glass, the pot is in a place where it could easily be knocked by a passer-by, and the bomb thereby detonated. There is a therefore a close world where the iron pot is broken. We can again exaggerate the case: the wealthy hater of sturdy things has organised a whole series of bombs and such like, any one of which could easily be detonated to blast the pot into pieces. In quite a number of close worlds the pot is broken. Since the iron pot is broken in at least a few close worlds, it is the case that, as Vetter understands

[3] The latter is ambiguous, and can be used to mean what the former does.

'easily', it could easily be broken, and so according to (F-V'), the iron pot is fragile. But that is false.

A response on Vetter's behalf might point out that the sturdy iron pot with a bomb and sensitive detonator is a mimic, a well-known counterexample to (F-S) and so to (SA) (see above). So the standard view fares equally badly. Nonetheless, on closer inspection, I think this case confirms the standard view. Yes, (F-S) is false. But why? Because the causal path from the stressing of the pot to its breaking is not of the right sort. In a fragile object, a light stress should lead directly to breaking, not via a bomb and the much greater stress it causes. The natural response to the mimic case is to focus on the nature of the stimulus. That focus makes sense on the standard account of dispositions. It is natural to ask: maybe the analysis can be reformulated to make the description of the stimulus more specific, so ruling out the bomb and detonator as suitable causal routes? On the other hand, Vetter's stimulus-free account of dispositions cannot make sense of this kind of interest in the stimulus. Correspondingly there isn't any avenue for amending or re-interpreting the details of (F-V') to avoid the problem of the iron pot. For the only machinery Vetter offers us concerns the number of close worlds in which the manifestation occurs. And as our examples show that's the wrong kind of machinery for this job.

4. *The problem of distinct dispositions*

The previous section gives a reason for thinking that we need stimuli to characterize dispositions. This section gives another reason. While Vetter has a response to this new reason, we shall see that her response exposes her view to the problem of non-conditional possibilities – the problem for the standard account that in my view gave Vetter's account a significant (but, we shall see, only prima facie) advantage.

A second reason to think that we need stimuli to individuate dispositions is the fact that distinct dispositions can have the same manifestation. Gravitational mass and electric charge both manifest themselves with a force. Vetter's manifestation-only approach would seem to require us to regard these as the same disposition. Clearly they are not. The standard view holds that they are distinguished by their different stimuli.

In fact Vetter has a means of dealing with this problem (although not explicitly advertised as such). She holds that for some dispositions the manifestation is itself conditional. In the case of charge the manifestation is a conditional of the form: if the object is at some distance from another charge, then it will

experience such-and-such a force. So in fact gravitational mass and charge are different dispositions because they do have different manifestations after all – different conditional manifestations.

One might reasonably complain that this does rather look as if stimuli are being smuggled back into the picture.[4] In broad-brush terms, the standard view is that dispositions are closely related to some conditional relationship between a stimulus condition and a manifestation condition. The detailed debates have been about the precise analysis of this relationship (or whether the relationship is analysable at all). In analyzing these dispositions in terms of a conditional relation between one condition and another, Vetter's account of these dispositions looks to be just a different version of the standard view.

But the difficulties that this response cause for Vetter's view are in fact rather deeper than this. The analysis of these dispositions in terms of conditional manifestations means, I shall argue, that Vetter's view now suffers from the same problem – the problem of non-conditional possibilities – that I said afflicted the standard view when used as the basis of an account of possibility.

Let us briefly recapitulate the problem of non-conditional possibilities for the standard account of dispositions, (SA). Taking properties with dispositional natures (powers, potencies) as elements of our fundamental ontology seems to offer a route to an account of modality. When allied with the standard account of dispositions, this approach encounters an obstacle: the dispositional property grounds the truth of a conditional possibility ('possibly, if the stimulus occurs, the manifestation occurs'). But what we want is also to ground non-conditional possibilities (such as 'possibly, the stimulus occurs' and 'possibly, the manifestation occurs'). For example, the glass possessing the property of fragility may ground the truth that were the glass struck, it would break. This in turn entails that it is possible that, if the glass is struck, it breaks. The latter is a conditional possibility. But we also want the non-conditional possibility, the possibility that the glass breaks. The conditional possibility, *possibly the glass breaks if struck*, does not entail the non-conditional possibility, *possibly the glass breaks*. For the former is consistent with the impossibility of the glass breaking, if it is impossible for the glass to be struck. So the disposition can only ground the non-conditional possibility, *possibly the glass breaks*, if a distinct non-conditional possibility, *possibly the glass is struck*, is also true. What grounds the latter? Since it is a non-conditional possibility, it cannot, for the reasons just given, be grounded in a dispositional property alone – it needs to be grounded in a dispositional property *plus* some other non-conditional possibility. The striking of the glass may be the manifestation of some other disposi-

[4] Vetter does present other, good arguments for her view, which I do not discuss here.

tion. But that other disposition cannot ground the non-conditional possibility, *possibly the glass is struck*, unless the stimulus for that other disposition is also a possible occurrence. So we get a regress. Not all non-conditional possibilities can be grounded in dispositions, if the standard account is correct. In a nutshell, the problem is this: we want to account for non-conditional possibilities. But because the standard account analyzes dispositions in terms of conditionals (relating stimuli and manifestations) the standard account can account only for conditional possibilities, not for non-conditional possibilities.

Vetter's approach, with dispositions characterized by manifestations alone, seemed to avoid this. If something is breakable (dispositional), then it can be broken (non-conditional). Dispositions are analyzed directly in terms of the possibility of the manifestation occurring. That is (so it would appear) a non-conditional possibility. So Vetter avoids the problem of non-conditional possibilities.

So it seemed. But that was before we considered the problem of distinct dispositions. Different dispositions can have the same non-conditional manifestation. Vetter's best response to this is to argue that in such cases the manifestation is not non-conditional after all. For properties such as charge, the manifestation is itself a conditional. In which case, such dispositions can only ground conditional possibilities. They do not ground non-conditional possibilities.

So the problem of non-conditional possibilities could well be a problem for Vetter's view, as well as for the standard view. Nonetheless, for all that has been said, the problem is not inevitable. For these properties (charge, mass) that have conditional manifestations are only some of the properties there are. Maybe some other properties have non-conditional manifestations. If so these can do the work of grounding enough of the relevant non-conditional modal truths. We might need only a few non-conditional possibilities – the conditional possibilities then take over, and generate further non-conditional possibilities (because the antecedents of the conditionals are satisfied, and so the consequents are satisfied also). For example, consider a large number of conditional propositions $Sx \rightarrow Mx$, $Mx \rightarrow Nx$, $Nx \rightarrow Ox$, $Ox \rightarrow Px$, $Px \rightarrow Qx$ and just one non-conditional proposition Sx. Then in addition to our one initial non-conditional possibility, $\Diamond Sx$, we get also the other non-conditional possibilities $\Diamond Mx$, $\Diamond Nx$, $\Diamond Px$, and $\Diamond Qx$.[5]

[5] Strictly, as noted above, to get the additional non-conditional possibilities, e.g. $\Diamond Mx$, we need not only one original non-conditional possibility, such as $\Diamond Sx$, and the conditional possibility, $\Diamond(Sx \rightarrow Mx)$, but we also need their compossibility, $\Diamond(Sx \,\&\, Sx \rightarrow Mx)$, or, better $(\Diamond Sx \,\&\, \Diamond(Sx \rightarrow Mx)) \rightarrow \Diamond(Sx \,\&\, Sx \rightarrow Mx)$. (The latter is preferable because it does not involve redundancy, as $\Diamond(Sx \,\&\, Sx \rightarrow Mx)$ does.) Either way, this introduces yet further modal truths for which, it would appear, Vetter's account cannot supply the grounds.

But that won't work because we don't have any *relevant* non-conditional modal truths. We started with the case of fragility (or, better, breakability). It looks as if, on Vetter's view, this grounds the possibility of breaking. So we get a non-conditional modal truth straight away. That, however, is misleading. First of all, it is far from clear that fragility/breakability is a genuine, sparse property at all. Perhaps it is a merely abundant property, and not part of any serious ontology. Even if we admit this property into our ontology, it is clearly not a fundamental property. And the nature, including modal nature, of non-fundamental properties supervenes on (or, is grounded in) the nature of the fundamental properties. What determines what is possible and not possible should be the fundamental properties alone. The existence of non-fundamental properties should not add any further possibilities not already determined by the fundamental properties.[6]

So let's focus on the fundamental natural properties. And let's assume (though this is far from assured) that charge is amongst these. Coulomb's law of electrostatic attraction tells us that the force, F, between two charges, q_1 and q_2, separated by a distance r is given by:

(C) $F = \epsilon q_1 q_2/r^2$.

Charge then is the property whose dispositional nature, according to Vetter (2015: 61), is given by the manifestation (where e is the charge on x):

(Q) $\forall r \forall q$: (x is at distance r from a charge of $q \rightarrow x$ exerts a force of $F = \epsilon eq/r^2$)

This manifestation is a conditional. We can expect all the fundamental properties to be like charge. For example, inertial mass is central to Newton's second law: $F = ma$. So a mass m is the dispositional property with manifestation:

(M) $\forall F$(x is subject to a force $F \rightarrow x$ experiences an acceleration of $a = F/m$).

Any dispositional property, then, whose nature is related to a law in this way, will have a conditional manifestation. This includes the fundamental properties that ground modality according to Vetter.

The preceding paragraph has argued that *all* the dispositional properties that ground modality have conditional manifestations. In which case they can-

[6] Consider this parallel (Bird 2016). Dispositional properties (properties with dispositional natures or essences) have been invoked to explain the laws of nature. Do we need non-fundamental properties to explain the existence of non-fundamental laws? No, non-fundamental laws supervene on the fundamental laws, and the latter are determined by the fundamental dispositional properties. So the fundamental properties suffice to fix the fundamental laws and the non-fundamental laws; we don't get any further non-fundamental laws from the non-fundamental properties. I suggest that the same reasoning applies to the grounding of modality in dispositional properties.

not ground non-conditional possibilities.[7] Vetter's view therefore also suffers from the problem of non-conditional possibilities, and so, despite initial appearances, it does not have the advantage, relative to the standard account of dispositions, of avoiding that problem.

5. *On linguistics and metaphysics*

Before concluding, I want to comment briefly on the role of arguments about the correct analysis of dispositional concepts in a debate about the metaphysics of modality. We started with a discussion concerning the correct analysis of dispositional concepts. And this was supposed to underpin a conclusion about the grounding of modality. Is there a legitimate connection between these?

Vetter (2014, 2015), drawing in part on work in linguistics by Kratzer (1981), presents a great deal of evidence that it is natural to regard certain dispositional expressions as characterized by manifestations alone and as expressing some kind of possibility. The examples used are expressions denoting properties of middle-sized goods and of people such as 'fragile' and 'irascible' or, from German, 'zerbrechlich' and 'erregbar'. That in turn informs the analysis of modality – possibility is grounded in dispositional properties, where 'dispositional' is understood as implicating a manifestation condition only, and not a stimulus condition.

I strongly suspect that the analysis of terms such as those just mentioned tells us very little about the metaphysics of modality. The hypothesis that important, general, and fundamental aspects of metaphysics, such modality and the laws of nature, are grounded in properties with a 'dispositional' nature is plausible. Its plausibility depends on the explanatory work it can do. It should not depend on the precise analysis of the dispositional terms we use in everyday life.[8] Why would we expect the semantics of 'fragile' or 'erregbar' to tell us anything about the metaphysical nature of fundamental properties such as charge or mass (or whatever the fundamental properties turn out to be)? When we say that the latter have a dispositional nature, that should not be held hostage to our discoveries about fragility. The problem is that Vetter's approach to modality looked most promising when viewed from the perspective of fragility and Erregbarkeit. What she says about charge is rather different. Yes, it

[7] It is important to note that I am not saying that there is something unsatisfactory about conditional possibilities (nor that properties with conditional manifestations are not real). Rather, I am saying that these possibilities cannot be all the possibilities that there are – a materially adequate account of possibility needs non-conditional possibilities as well as conditional possibilities.

[8] This comment does not do justice to the highly sophisticated account of potentiality and then of possibility that Vetter gives us. Still, I do think that the basic problem I identify remains.

looks as if it fits the same (stimulus-free) mould, because there is still only a manifestation. But now the manifestation is a conditional. And that raises the problem of non-conditional possibilities. That problem appeared to be absent only when we looked at properties such as fragility and Erregbarkeit – which, in my view, are irrelevant to the metaphysics of modality.

This kind of move – drawing conclusions about fundamental aspects of modality from evidence regarding our use of everyday expressions – is widespread in this area, and so this concern applies well beyond Vetter's work. We need to think carefully about whether this is a legitimate inference and if so why.

6. *Conclusion*

I think that the problem of non-conditional possibilities is an important and general one. We want to tie the nature of fundamental properties to their explanatory role. That is, the nature of such properties is such that they generate the laws of nature. But the laws of nature are conditional in nature. So the characterization of the fundamental properties will be conditional too – that's the case whether we think of this as a conditional relationship between a stimulus and a manifestation (standard account) or as a conditional embedded entirely within the manifestation (Vetter's account). But if our supervenience/grounding base is conditional in nature, it is difficult to see how this base can ground the non-conditional possibilities.

Alexander Bird
alexander.bird@kcl.ac.uk
Peter Sowerby Professor of Philosophy and Medicine, King's College London

References

Bird, A., 1998, "Dispositions and antidotes", in *Philosophical Quarterly*, 48: 227-234.

Bird, A., 2007, *Nature's Metaphysics: Laws and properties*, Oxford University Press, Oxford.

Bird, A., 2016, "Overpowering: How the powers ontology has over-reached itself", in *Mind*, 125: 341-383.

Borghini, A. and Williams, N.E., 2008, "A dispositional theory of possibility", in *Dialectica*, 62: 21-41.

Jacobs, J.D., 2010, "A powers theory of modality: Or, how I learned to stop worrying and reject possible worlds", in *Philosophical Studies*, 151: 227-248.

Johnston, M., 1992, "How to speak of the colors", in *Philosophical Studies*, 68: 221-263.

Kratzer, A., 1981, "The notional category of modality", in H.-J. Eikmeyer, H. Rieser

(eds.), *Words, Worlds, and Contexts: New approaches in word semantics*, De Gruyter, Berlin: 38-47.

Lewis, D.K., 1997, "Finkish dispositions", in *Philosophical Quarterly*, 47: 143-158.

Manley, D. and Wasserman, R., 2008, "On linking dispositions and conditionals", in *Mind*, 117: 59-84.

Martin, C.B., 1994, "Dispositions and conditionals", in *Philosophical Quarterly*, 44: 1-8.

Vetter, B., 2014, "Dispositions without conditionals", in *Mind*, 123: 129-156.

Vetter, B., 2015, *Potentiality: From dispositions to modality*, Oxford University Press, Oxford.

A strange kind of power: Vetter on the formal adequacy of dispositionalism

David Yates

Abstract: According to dispositionalism about modality, a proposition <*p*> is possible just in case something has, or some things have, a power or disposition for its truth; and <*p*> is necessary just in case nothing has a power for its falsity. But are there enough powers to go around? In Yates (2015) I argued that in the case of mathematical truths such as <2+2=4>, nothing has the power to bring about their falsity or their truth, which means they come out both necessary and not possible. Combining this with axiom (T): $p \supset \Diamond p$, it is easy to derive a contradiction. I suggested that dispositionalists ought to retreat a little and say that <*p*> is possible just in case either *p*, or there is a power to bring it about that *p*, grounding the possibility of mathematical propositions in their truth rather than in powers. Vetter's (2015) has the resources to provide a response to my argument, and in her (2018) she explicitly addresses it by arguing for a plenitude of powers, based on the idea that dispositions come in degrees, with necessary properties a limiting case of dispositionality. On this view there is a power for <2+2=4>, without there being a power to bring about its truth. In this paper I argue that Vetter's case for plenitude does not work. However, I suggest, if we are prepared to accept metaphysical causation, a case can be made that there is indeed a power for <2+2=4>.

Keywords: modal dispositionalism; formal adequacy; powers; truthmaking; grounding.

1. *Dispositionalism and formal adequacy*

According to dispositional theories of modality (hereafter 'dispositionalism'), a proposition <*p*> is possible just in case something has (or some things have) a power for <*p*>, and since necessity is the dual of possibility, <*p*> is necessary just in case nothing has a power for its falsity.[1] Depending on how we flesh out the notion of a power for <*p*>, dispositionalist theories with very different metaphysical and formal properties will result. In this paper I shall discuss Vetter's dispositionalism, paying particular attention to her arguments

[1] Versions of dispositionalism are defended in Williams & Borghini (2008); Contessa (2010); Jacobs (2010); and Vetter (2015), which provides by far the most detailed and sustained defence, addressing not only the material adequacy of dispositionalism, but also the formal adequacy of dispositionalist modal logic. It should be noted that for present purposes, I shall not distinguish powers from dispositions, and use the terms interchangeably.

ISSN (print) 2281-8618-ETS DOI: 10.4454/philinq.v8i1.277

that there are enough powers available to yield a formally adequate dispositionalist modal logic. I shall argue that Vetter has not made a compelling case for the powers required to do the job, and suggest an alternative argument on behalf of the dispositionalist based on the notion of metaphysical causation, and the corresponding idea of metaphysical causal powers. More on that presently. First, let's clarify the central notion of a power for <*p*>:

> Metaphysical modality is puzzling because it does not fit into the schema of objects-with-properties. It seems to consist of facts that float free of any particular object: its being possible that there are talking donkeys, for instance. One of the attractions of Lewisian modal realism is that it anchors those free-floating facts in objects. One of its drawbacks is that the objects are otherworldly donkeys for or against whose existence we can in principle have no evidence. Dispositionalism promises to share the attraction without succumbing to the drawback: it, too, anchors possibilities in objects. But its objects are just the ordinary objects of this, the actual, world, with which we are in regular epistemic contact…By anchoring them in the dispositions of such objects, dispositionalism promises a plausible story about the epistemology of modality (Vetter 2015: 11).

On the basis of the quoted passage, it seems initially that Vetter is concerned solely with what I shall herein refer to as *efficient causal powers*: the powers of objects to bring about or cause events. I adopt a Kimian conception of events for present purposes,[2] where events are understood as objects having properties at times, and hence not distinguished from states. So conceived, efficient causal powers will often be powers to bring about change, but need not be. An object could have the power to keep moving in a straight line, for instance, or to hold its shape.[3] I am not committed to the view that efficient causal powers are exclusively powers to cause physical events, or exclusively due to physical properties. What is important is that efficient causal powers are powers to cause states, or changes in the states, of objects – to cause them to have or continue having certain properties. Efficient causal powers manifest when they are appropriately stimulated, or alternatively, when they meet their reciprocal partner powers,[4] and are *individuated* by these causal roles. Negative charge, assuming it to be an efficient causal power, is partially individuated by the fact that its bearers exert repulsive forces on each other – having this causal role is part of what makes negative charge the property it is. This claim is typically unpacked in terms of relational individuation: efficient causal powers are individuated by their places

[2] See Kim (1976).

[3] Changes can then be understood in terms of sequences of Kim events.

[4] See Bird 2007a for the stimulus-manifestation model, and Martin 2008 for the partner powers model.

in a second order causal structure, composed of powers (and possibly also non-powers) standing in primitive stimulus-manifestation relations.[5] And we can in principle come to know this structure through scientific investigation.

What is it to have an efficient causal power for <*p*>? I shall say that an efficient causal power for a proposition <*p*> is a power to cause a truthmaker for <*p*>, and thereby *bring about* that <*p*> is true. Call the position that results from combining dispositionalism with the claim that a power for <*p*> is an efficient causal power to bring it about that *p*, *narrow dispositionalism*. Narrow dispositionalism explains the epistemology of modality because (at least in principle) we have epistemic access to the defining roles of efficient causal powers. Could I have been an astronaut? Yes, provided there's a way of combining the actual efficient causal powers such that had they been so combined, their manifestation would have been, or included, my being an astronaut. Judging the truth values of such claims may be difficult, but it is at least based on our knowledge of the world, and of the efficient causal powers that are instantiated here.

In an earlier work (Yates 2015) I argued that narrow dispositionalism faces problems of both material and formal adequacy, as follows. Obviously <2+2=4> is possible, but there doesn't seem to be anything with the power to *bring it about* that 2+2=4. There is no efficient causal power with a truthmaker for <2+2=4> as its manifestation, because the truthmakers for mathematical propositions are beyond the reach of efficient causation. Narrow dispositionalism thus entails that <2+2=4> is not possible, which is a serious problem of material adequacy. However, it also (correctly) entails that <2+2=4> is *necessary*, since there is no efficient causal power to bring about that <2+2=4> is false either. As Vetter notes (2018),[6] efficient causal powers are *symmetric* with respect to necessary propositions and their negations, in that they are not able to bring about either. As a necessary condition on being a theory of metaphysical modality, dispositionalism ought to yield a modal logic satisfying axiom T: $\Box p \supset p$ (equivalently, $p \supset \Diamond p$). But given (T), $\neg\Diamond p$ and $\Box p$ are inconsistent. From (T) and $\Box p$ we can conclude *p*, but from (T) and $\neg\Diamond p$ we can conclude $\neg p$. Hence narrow dispositionalism either: (i) results in an inconsistent modal logic, or (ii) is not a theory of metaphysical modality.

In response to this problem, I suggested weakening narrow dispositionalism, in a manner suggested by Pruss (2011) on independent grounds, so that <*p*> is possible just in case either: (i) <*p*> is true, or (ii) something has (or some

[5] See Bird 2007b for a defence of the relational individuation in pure powers ontologies; and see Yates 2018 for a defence of the relational individuation of powers in a structure that also includes realized non-powers.

[6] Note that Vetter doesn't use the term 'efficient causal power', but she is clearly referring to powers to cause events or changes, which is how I am using the term in question in the present work.

things have) an efficient causal power for <*p*>. Powers take care of the possibility of propositions that are possible but false, with the possibility of true propositions such as <2+2=4> grounded in their truth, rather than in powers to bring about their truth. Necessary propositions, correspondingly, are those truths whose truth value nothing has the power to change. Vetter, by contrast, argues that the problem with narrow dispositionalism lies not in its focus on powers, but in its focus on *efficient causal* powers. By way of response, she argues that the version of dispositionalism she had earlier defended has the resources to ground all modal truths in powers, resulting in a more elegant and unified theory.[7] Vetter's solution is to allow powers that are not efficient causal, and which necessarily always manifest, without stimuli or partner powers. I considered but dismissed this strategy on the grounds that the powers in question would be so far removed from efficient causal powers as to be powers in name only, posited ad hoc in order to solve the formal adequacy problem; Vetter responds by giving independent arguments for believing in them.

My solution, then, was to embrace what I shall herein refer to as *disjunctive dispositionalism*, with efficient causal powers taking care only of unactualized possibilities, while Vetter prefers a version of dispositionalism according which some powers are not efficient causal powers, and the resulting *plenitude of powers* does all the work of grounding modality. I shall refer to Vetter's theory as *broad dispositionalism*, because it recognises a broader variety of powers than just powers to cause events, which I had earlier tacitly assumed to be all the powers that there are. Broad dispositionalism seems to be a more elegant and unified theory, since it explains all modal truths in terms of primitively modal powers. By contrast, disjunctive dispositionalism grounds some possible truths in powers, and others solely in their truth.[8]

The rest of this paper proceeds as follows. In §2 I argue that Vetter's case for plenitude fails, and hence that broad dispositionalists need an alternative argument for the missing powers. In §3 I offer an alternative argument for plenitude based on the idea that if truthmaking is a form of metaphysical causation, then truthmakers can plausibly be said to have the power to make true. Whether the resulting form of broad dispositionalism is any more unified than disjunctive dispositionalism depends, I suggest, on whether such powers are deservedly so-called.

[7] In Vetter 2018 she argues that the dispositionalist theory detailed in Vetter 2015 has the resources to respond to the arguments I gave in Yates 2015.

[8] Here is Vetter on disjunctive vs. broad dispositionalism: "I think that [disjunctive] dispositionalism is an acceptable last resort for the dispositionalist […] I take it, however, that [broad] dispositionalism […] is preferable if it is to be had: it gives a unified dispositional picture of metaphysical modality", (2018: 4 (ms)).

2. *Vetter's case for plenitude*

Vetter lays much of the groundwork for her case for plenitude in her (2015), but develops and clarifies the case in her (2018) response to my (2015), so I shall focus primarily on the latter work in this section. There are several reasons why one might want to deny that anything has a power for <2+2+4>. As Vetter notes (2018),[9] such a power: (i) could not possibly fail to manifest, at any time; (ii) would have neither a stimulus condition nor reciprocal partner powers; (iii) would not be a power to bring about what it was a power for, i.e. that 2+2=4. It follows right away that a power for <2+2=4> could not be an efficient causal power, and if one thought, as I had previously supposed, that all powers are efficient causal, then it would follow in addition that there is no power *simpliciter* for <2+2=4>. There are two possibilities, in principle, for powers that are not efficient causal: (A) noncausal powers, and (B) causal powers whose manifestations are not events. I will later suggest a strategy based on powers of type (B), but for now I shall focus on Vetter's arguments, which aim to defend powers of type (A).

Vetter's defence of plenitude comes in two parts. In the first, she gives an *argument from degrees* for the claim that if *x* has an intrinsic property P necessarily, then *x* is maximally disposed to P.[10] Although Vetter doesn't explicitly frame it as such, the argument, if successful, will establish that there are powers for necessary properties of either type (A) or type (B). The idea that things have powers for their necessary intrinsic properties is bound to take us beyond the range of efficient causal powers in *some* way. In the second part, she defends the argument from degrees against a potential objection that necessary properties are not suitable disposition manifestations, by arguing that there are independently motivated cases of dispositions that are relevantly similar to those that the argument from degrees is intended to establish. At this point, Vetter makes clear that she takes powers for necessary properties to be of type (A). I address the argument from degrees in this section, and return to the issue of noncausal powers in §3.

The argument from degrees depends on the following proportionality principle: where P and Q are contradictories, the degrees of *x*'s dispositions to P and Q are inversely proportional. Vetter uses 'contradictory' in a fairly informal sense here, sometimes referring to contraries, and sometimes to logically contradictory predicates. Examples given include: break vs. remain unbroken, remain calm vs. get angry, and talk vs. remain quiet. As the disposition of a

[9] In noting this, Vetter is agreeing with my claims (Yates 2015) about the nature of putative powers for propositions such as <2+2=4>, but she goes on to argue that there is after all good reason to posit them.

[10] See Vetter 2018: §2 for a succinct presentation; for the full details, see Vetter 2015: chs. 2-3).

vase to break goes up, so its disposition to remain unbroken goes down; the more one is disposed to get angry, the less one is disposed to remain calm; and the more one is disposed to talk, the less one is disposed to remain quiet. I shall focus on contradictories rather than contraries. Vetter's proportionality principle entails that at least one of '*x* is disposed to P' and '*x* is disposed to Q' is true. I shall refer to this latter claim as *universality*. Given universality, it's reasonable to conclude that where *x* is necessarily P, *x* is maximally disposed to P and minimally (i.e. not at all) disposed to Q. Here's how Vetter puts it:

> As a limiting case of the proportionality principle, we can say that if M'ing and N'ing are contradictories (and both qualify as genuine properties of *x*), as *x*'s disposition to M reaches the maximum, its disposition to N must reach a minimum, and vice versa; that is, *x* is maximally disposed to M just in case *x* is minimally disposed to N. The absolute minimum of a disposition, in turn, should be the lack of it. So we can say that *x* is maximally disposed to M iff *x* is not at all disposed to N. For any pair of contradictory predicates M and N, then, it follows that at least one of '*x* is disposed to M' and '*x* is disposed to N' must be true. Applied to the cases that interest us, this means that I must have either a disposition to be human or a disposition not to be human; either a disposition to be dancing-or-not-dancing, or a disposition not to be dancing or- not-dancing. Faced with the alternative, I take it, we should prefer the necessarily always manifested disposition to the necessarily never manifested disposition (Vetter 2018: 8 (ms)).[11]

By proportionality, whatever it is that has the property of being such that 2+2=4 intrinsically has inversely proportional degrees of the power for <2+2=4>, and the power for its negation. Given that <2+2=4> is necessary, we should say that the entities in question are maximally disposed for <2+2=4>, and not at all disposed for its negation. What kind of entities are they? Vetter is clear that it is the fundamental truthmakers for <2+2=4>, whatever they might be, that have the property of being an *x* (or some *x*s) such that 2+2=4 intrinsically, so it's these entities that will come out as being maximally disposed for <2+2=4>. These object could be sets, numbers conceived as abstract objects, or concrete particulars. For present purposes I needn't commit to a position on the nature of the truthmakers for arithmetic propositions, but I will assume in what follows that they are Platonic numbers for ease of exposition. More important is to note that Vetter takes the bearers of the power for <2+2=4> to be its *truthmakers*:

> If numbers are bona fide abstract objects, and the truth of <2+2=4> is a matter of such abstract objects standing in certain relations, then those are the objects whose

[11] By "genuine properties" here, Vetter means *intrinsic* properties.

powers we should look to….The powers that ground the possibility that 2+2=4 will be powers to have just those features in virtue of which the relevant objects ground the truth of <2+2=4> (Vetter 2018: 19 (ms)).

So, numbers have certain features in virtue of which they make propositions such as <2+2=4> true. They have these features intrinsically and necessarily, and so – assuming the argument from degrees works – are maximally disposed to have them. In other words, the power for <2+2=4> is the power to have the features in virtue of which its truthmakers make it true.[12]

Before proceeding, let us first clarify the parenthetical constraint in the proportionality principle, which is apt to be a little confusing: "and both qualify as genuine properties of *x*" could be taken to mean that *x* actually has both of the properties expressed by the predicates in question, but that is a contradiction. The predicates P and Q do not express dispositional properties here, but their manifestations, and given that they are by hypothesis contradictory, there is no way *x* could have both properties at once: one cannot both break and not break, or remain calm and get angry. There's no contradiction in the supposition that *x* is (to inversely proportional degrees) both *disposed* to break, and *disposed* to not break, but again, P and Q don't refer to dispositions. Rather, they will refer in this case to *breaking*, and to *not-breaking*. Hence, the intrinsicality condition is intended to restrict quantification to properties that are intrinsic to their bearers *when instantiated*.

Let us now clarify the relationship between proportionality and universality. Proportionality is the claim that for any pair of contradictory predicates P and Q expressing intrinsic properties of *x*, the degrees of *x*'s dispositions to P and Q are inversely proportional. Universality is the claim for any such pair of predicates, at least one of '*x* is disposed to P' and '*x* is disposed to Q' is true. Why does the latter follow from the former? Here Vetter assumes that the degrees of dispositions can be quantified, and that being X% disposed to Φ involves Φ-ing in X% of cases. It follows, since P and Q are contradictories, that their degrees must always add to 100%, because if I P in X% of cases, I must not-P in the remaining (100-X)% of cases. From this it follows right away that *x* is maximally disposed to P iff *x* is not at all disposed to Q. Hence, from the non-instantiation of one of a pair of contradictory dispositions, we can conclude the maximum degree of the other. It is therefore not possible to be neither disposed to P, nor disposed to Q. This is a somewhat surprising claim – who knew that things had so many powers? Granting the inference from proportionality to universality, I shall now argue that there are counterexamples to the latter, hence to the former.

[12] I shall return to the relationship between truthmaking and powers in §3.

Vetter's proportionality principle is the claim that where P and Q express contradictory intrinsic properties, like the properties of singing and of not singing, *x*'s *dispositions* to P and to Q are inversely proportional, from which it follows, as we have just seen, that at least one of the corresponding disposition ascriptions is true. This all sounds plausible when we consider simple dispositions like the disposition to sing or the disposition to break: all actual things are to some degree disposed to break, and to an inversely proportional degree disposed to not break. If we allow the degree of one of these dispositions to go to zero, the other goes to maximum, such that the complete absence of either is sufficient for the presence of the other, and nothing can fail to have both. Crucially, however, Vetter considers only simple dispositional idioms constructed by prefixing a predicate expressing an intrinsic property, such as the property of singing, with 'the disposition to'. I am either singing or not singing; I can't be both, and I can't be neither. It's plausible that the more often I sing, the more *disposed* I am to sing, and the less often, the less so disposed. Someone who never sings is not at all disposed to sing, but then such a person must be maximally disposed to not sing. Hence, as per universality, everyone is (to some degree) either disposed to sing, or disposed to not sing. For simple dispositions such as these, it is difficult to avoid Vetter's conclusion. However, when we attribute dispositions, we often refer not only to the their manifestation properties, but also to the *conditions* under which they manifest. I shall now argue that in such cases, it is possible to have neither of the dispositions in question to *any* degree.

Disposition ascriptions of the form '*x* is disposed to M when C' often imply causality. Consider, for instance, the disposition to sing *when it rains*. Here 'when' implies 'when and because', and does not merely express a correlation between rain and singing. The manifestation property is the same as in the disposition to sing, but the specification of the disposition now refers to a stimulus condition. The disposition to sing when it rains is distinct from the disposition to sing when the sun shines, and both are distinct from the disposition to sing, *simpliciter*. Vetter's proportionality principle refers only to manifestation properties, so it implies that the disposition to sing when it rains and the disposition to not sing when it rains are inversely proportional. I don't think they are. Suppose for the sake of argument that the rain makes no difference to the probability of my singing. Suppose further that I am to some degree disposed to sing, *simpliciter*. Now suppose that it's raining on some occasion, and I sing. Is this a manifestation of the disposition to sing *when it rains*? It can't be, on the assumption that the latter disposition implies causality. The rain made no difference to my singing on this occasion, I just happened to manifest my disposition to sing *simpliciter*, while it was raining.

If the rain makes no difference to whether or not I am singing, then I lack the disposition to sing when it rains. By parity of reasoning, I also lack the disposition to not sing when it rains. If universality is false for this disposition pair, then so is proportionality: the absolute minimum of each disposition is consistent with the absolute minimum of the other. We can see that proportionality fails here without considering just the minima. Imagine that as a young romantic, I was strongly disposed to sing when it rained. Now, bored with life, the rain, and everything, I do so only occasionally; but I am not – at least not yet – so jaded that the rain ever *prevents* me from singing. If these arguments are correct, then a reduction in the degree of a disposition does not imply an increase in the degree of its contradictory, so it's false that the degrees of contradictory dispositions must add to 100%.

It's because the 'when' in many common disposition ascriptions implies causality that proportionality fails. I am certainly sometimes disposed to sing *while* it rains, but there's no causal connection, so I'm not disposed to sing *when* it rains; nor, for the same reason, do I have the contradictory disposition to not sing when it rains. It's intuitive, however, that contradictory 'when' dispositions must instantiate *some* form of proportionality relationship. What then is the relationship? Let's say that being X% disposed to Φ when C involves Φ-ing when (and because) C in X% of cases. According to Vetter's proportionality principle, the degrees of contradictory dispositions must add to 100%, so that (for instance) being 60% disposed to Φ when C entails being 40% disposed to not-Φ when C. If the preceding arguments are correct, this cannot be right, because I can be 0% disposed to sing when it rains, and 0% disposed to not sing when it rains. However, there is a weaker proportionality relationship that does hold. Suppose I am 60% disposed to sing when it rains. What follows that I am *at most* 40% disposed to not sing when it rains. The degree of a disposition places an upper limit on the degree of its contradictory, but crucially, does not determine a *lower* limit. That I am 60% disposed to sing when it rains does not entail that I am to *any* degree disposed to not sing when it rains, since it doesn't guarantee that the rain ever prevents my singing. The degrees of two contradictory dispositions add to *at most* 100%, but to at least 0%. And one who doubts the existence of a power for <2+2=4> might well suspect that this is a case where the sum of the degrees is 0%: nothing has, to any degree, either a power for <2+2=4>, or a power for its negation.

Now Vetter might offer the following response to this objection, based on her independently motivated view that dispositions are individuated by their manifestations alone.[13] According to Vetter, disposition ascriptions that seem

[13] Vetter 2014; 2015. This is one of the core commitments of Vetter's (2015) account of disposi-

to be individuated by both a stimulus and a manifestation property, such as the disposition to sing when it rains, should be understood as dispositions individuated solely by a complex manifestation. In the present case, this would be the disposition to be caused to sing by rain, whose manifestation is the complex property of *being caused to sing by rain*. Strictly speaking, this is not an intrinsic manifestation property, and so is outside the scope of Vetter's proportionality principle. However, let us assume that it is a suitable manifestation property for a disposition, and relax the "genuine property" condition to allow such properties in the scope of the principle. Our question now is this: does universality hold for dispositions so understood?

Whether we think of dispositions as having separate stimuli and manifestation properties or complex manifestation properties significantly affects how we should go about constructing their contradictories. If we think of dispositions as individuated by both a stimulus and a manifestation property, then for a disposition of the form 'the disposition to M when C', it's easy to construct contradictories: negate the manifestation predicate to get 'the disposition to not-M when C', leaving the stimulus term unchanged. The disposition is then contradictory in the sense that it is the disposition to do the opposite thing under the same conditions. By contrast, in the case of dispositions individuated solely by a complex manifestation, of the form 'the disposition to be caused to M by C', it's harder to construct the contradictory, because there are now two places where the negation can be introduced. If we construct the contradictory of this disposition by wide scope negation over its complex manifestation, then the contradictory will be: the disposition *not* to be caused to M by C, but if we use narrow scope negation, the contradictory will be: the disposition to be caused to *not*-M by C.

In the present case, wide scope negation yields: the disposition not to be caused to sing by rain. Plausibly, everyone either has the disposition to be caused to sing, or the disposition not to be caused to sing, by rain, so universality is safe. But is wide scope negation the *right* way to generate contradictories of dispositions as Vetter conceives them? Consider how the disposition not to be caused to sing by rain should be rendered in the commonplace 'when' parlance. If I am disposed not to be caused to sing by rain, it clearly doesn't follow that rain will *cause me to not sing.* Compare: a vase is disposed not to be broken by a fly landing on it, but it doesn't follow from this that a fly landing on it will cause the vase to *not* break. Hence the most natural way to render '*x* is disposed not to be caused to sing by rain' in the stimulus-manifestation

tions, but it is not clear to me whether or not she takes the argument from degrees to depend on it. Vetter's (2018) summary of the argument from degrees does not appeal to this commitment.

idiom is: not [*x* is disposed to sing when it rains]. But that's just the negation of an ascription of the disposition to sing when it rains, and not an attribution of the contradictory *disposition*. If contradictory dispositions could be generated merely by the negation of disposition ascriptions, then universality would be a logical truth: either *x* is disposed to Φ, or not [*x* is disposed to Φ].

In stimulus-manifestation terms, the contradictory of the disposition to sing when it rains is the disposition to *not* sing when it rains, and this seems more naturally rendered in Vetter's complex manifestation idiom by means of *narrow* scope negation: being caused to *not* sing by rain. So now it seems our pair of contradictories should be the dispositions: (i) to be caused to sing by rain, and (ii) to be caused to not sing by rain. But now we are right back where we started: it's possible to lack *both* of these dispositions, so once again universality is threatened, and with it proportionality. Even setting aside the above concerns about how best to render 'when' dispositions in terms of Vetter's complex manifestation idiom, it seems clear that (i) and (ii) are a pair of contradictory dispositions whose minima are mutually consistent.

Let's recap. Vetter's central claim is that where P is an intrinsic property, if *x* is necessarily P then *x* is maximally disposed to P. In order to establish this claim, she relies on *proportionality*: the claim that for any pair of contradictory predicates P and Q expressing intrinsic properties, *x*'s dispositions to P and Q are inversely proportional. The complete absence of one disposition then corresponds to the maximum degree of the other, and we can conclude, where *x* is necessarily P, that *x* is maximally disposed to P. However, there seem to be clear cases of disposition pairs, such as singing vs. not singing when it rains, whose minima are mutually consistent, and which therefore do not satisfy Vetter's proportionality principle. We can certainly set an upper limit of 100% on the sum of the degrees of two such dispositions, but the lower limit is 0%. Problematically, that allows a sceptic about plenitude to say that nothing is to any degree disposed for either <2+2=4> or its negation. I will now set aside the argument from degrees, and turn to Vetter's defence of the claim that necessary properties such as the intrinsic properties of abstract objects are suitable power manifestations.

3. *Efficient causal, noncausal, and metaphysical powers*

Vetter worries that the restriction to intrinsic properties in the argument from degrees may not be strict enough, and so addresses a potential objection that the features in virtue of which numbers make <2+2=4> true aren't the right kind of properties to be power manifestations. Now the argument given in §2 targets the proportionality principle, and applies to properties that are *clearly*

suitable to be disposition manifestations, such as the property of singing. Even if the properties of numbers are suitable power manifestations, I don't think Vetter's argument from degrees would suffice to show that anything has the relevant power. In this section, I shall focus on Vetter's defence of the claim that the properties of numbers are suitable power manifestations. Vetter argues by analogy that such properties could be the manifestations of necessarily manifesting noncausal powers. I don't think the analogy works, but it is nonetheless instructive: it tells us something about the kind of powers we need to ground the possibility of mathematical propositions, and thereby points the way to an alternative to the argument from degrees as a means of defending plenitude.

An efficient causal power *for* a proposition <*p*> is a power to bring about that <*p*> has a truthmaker. An efficient causal power for <the window is broken> is a power to break the window – to bring it about that <the window is broken> is true. But numbers – again assuming mathematical Platonism – just *are* the truthmakers of <2+2=4>. Furthermore, *qua* abstract objects, they could not possibly be the bearers of efficient causal powers. Indeed, that is arguably part of *what it is* to be abstract, rather than concrete. Since they are in Vetter's view powers for the features in virtue of which numbers make <2+2=4> true, their manifestation is just the instantiation of those features. Given that truthmakers are sufficient for truth, the manifestation of these powers in a sense results in the truth of <2+2=4>, but by no means brings it about. Powers for mathematical propositions, if they exist, are a strange kind of power. Vetter's strategy in defending them is to argue that for every strange feature they have, there is an independently motivated case of a power with that feature. The strange features Vetter considers are: (i) necessary manifestation properties, (ii) lack of stimulus conditions or partner powers, (iii) not being causally related to their manifestations. Having argued for powers that have some of these features individually, Vetter then challenges her objector to say why it should be any more problematic to have all of them.[14]

Although Vetter's purpose is to defend the argument from degrees by highlighting similarities between efficient causal and mathematical powers, I think the cases she gives actually highlight an important difference between them. As Vetter notes, there are some fairly straightforward cases of powers that manifest without conditions, such as the power of a massive object to curve spacetime, or the power of an electron to generate an electromagnetic field. It is therefore no problem for mathematical powers that they are condition-free. I also grant that there may be powers that manifest necessarily. Suppose, for example, that electric charge is the power to generate an electromagnetic

[14] Vetter 2018 §3. I grant here that the conjunction is unproblematic if each conjunct is.

field conditionally on being instantiated. Charge would in that case be a power with no stimulus conditions or partner powers, and it might also provide an example of a power that manifests necessarily. We would need to offer an argument that nothing could possibly interfere with its manifestation, but I will not attempt to do so here, since what I want to focus on is Vetter's defence of the claim that *noncausal* powers are unproblematic.

Why suppose there to be noncausal powers? The examples upon which Vetter draws are due to Nolan (2017). A volcano is disposed to smoke prior to eruption, but its stimulus, the eruption, happens after its manifestation, the smoking. This looks to be a genuine disposition ascription, but it's not clear how the disposition in question could be among the causes of its manifestation, for the disposition is triggered only after it has already manifested. It is also plausibly true that the Nile is disposed to flood after Sirius rises at or just before sunrise, but there is clearly no causal connection between Sirius' rising at dawn and the Nile's flooding. The disposition is triggered by a condition that could not possibly be a cause of its manifestation, so once again there is no obvious causal role for the disposition itself in relation to its manifestation.

Now at least in the cases mentioned above, it's natural to say that the objects in question have causal powers that ground the noncausal powers in question. A volcano's smoking and its eruption are caused by magma manifesting its power to flow upwards through weak points in the Earth's crust, and this is obviously an efficient causal power. Similarly, the Nile has an efficient causal power to flood when certain weather conditions obtain, and those conditions happen to correlate with the rising of Sirius at or before sunrise. This is not to deny that the relevant noncausal powers exist, and I'm happy to grant that they do. However, they are not *fundamental* noncausal powers, but grounded dispositional states that can be explained in terms of more basic efficient causal powers.[15] It's not clear that all noncausal powers are like this, and Nolan offers further examples that don't seem to be, for instance the dispositions of one of a pair of quantum entangled electrons to be spin up if the other is measured to be spin down.

Causal interpretations of what happens during measurement are problematic, because the causal influence would have to travel faster than light, which makes the relevant disposition look noncausal. And in this case, it's not obvious that there are any efficient causal powers available that could explain the correlation, which makes it look *fundamentally* noncausal. However, the interpretation of quantum non-locality is hugely controversial. Depending on what we say about causation, we might have no choice but to say that the mea-

[15] Nolan acknowledges this point, and anticipates that it may be argued that noncausal powers are not fundamental but grounded in causal powers; see Nolan 2015: 429-431.

surement of one entangled particle causes changes in the quantum state of the other.[16] Or it may be that quantum nonlocality is to be explained by local causality in a more fundamental structure such as the wavefunction on 3N-dimensional configuration space, with the fundamental dynamics giving rise to the appearance of nonlocality when projected onto our 3-dimensional space (Ismael & Schaffer 2016). But in that case, quantum entanglement seems similar to Nolan's more mundane cases after all, in which noncausal powers are grounded in efficient causal powers.

What I want to suggest now is that the gulf between the noncausal powers Vetter draws attention to, and putative mathematical powers, is much wider than she acknowledges. In fact, a power for <2+2=4> would seem to be so strongly noncausal that we might wonder why it deserves to be called a power at all. The disposition to smoke before erupting is noncausal in that the disposition itself is not among the causes of its manifestation. However, its manifestation property – smoking – certainly *has* causes, and its causes are among the grounds of the corresponding noncausal disposition. The volcano has the efficient causal power to emit smoke, and the efficient causal power to erupt, and it normally manifests these powers sequentially, which explains why it has the *noncausal* power to smoke before erupting. It's one thing, however, to identify a disposition that doesn't cause its manifestation; it's another thing entirely to posit a disposition whose manifestation doesn't *have* a cause.

When we ascribe noncausal dispositions like the disposition to smoke before erupting, we say something about the event-causal structure of the world, and it's that structure that makes the disposition-ascription true. The putative power for <2+2=4> is dissimilar in a crucial respect. If there were such a power, it could not be grounded in an efficient causal power for <2+2=4>, because there is no such power. The power for <2+2=4> would be a power for a proposition whose truthmaker is beyond the reach of efficient causation, and that's a significant *dis*analogy between it and the independently motivated cases of noncausal powers to which Vetter appeals. In sum, noncausal dispositions that are grounded in an underlying efficient causal structure give us no reason to believe that the necessary properties of numbers are suitable power manifestations. I grant that there are powers that don't cause their manifestations, but we need something stronger than that: powers whose manifestations are not part of the event-causal structure of the world at all.

Let's recap. Given that there are no efficient causal powers for propositions such as <2+2=4>, the broad dispositionalist has two options: either (A)

[16] See for instance Bigaj 2017. Note that the resulting superluminal causation is not to be confused with superluminal *signalling*, which is widely accepted to be ruled out by special relativity.

noncausal powers, or (B) powers that are causal but whose manifestations are not events. Vetter argues that there are powers for necessary propositions by arguing that there are powers for necessary intrinsic properties, as limiting cases of dispositionality, which comes in degrees. I have already expressed my doubts about that argument, but even if it works, it remains unclear that properties of numbers are suitable power manifestations, and so unclear that they are within the scope of the argument in the first place. Vetter argues that they are suitable power manifestations by analogy with other powers that manifest necessarily and non-causally. As we've seen, however, the sort of noncausal powers for which we have independent motivation seem to be grounded in efficient causal powers. But it's their total disconnection with efficient causation that makes the putative powers for propositions like <2+2=4> look like non-powers in the first place, so drawing attention to noncausal powers that are grounded in efficient causal powers doesn't help. In what follows I will suggest an argument for powers of type (B): powers that are in a sense causal, but which don't bring anything about, and whose manifestations are not events. This argument is an alternative to Vetter's argument for plenitude. It does not aim to establish that the necessary intrinsic properties in virtue of which numbers make <2+2=4> true are suitable power manifestations. Instead it aims to show that true propositions are themselves the manifestations of a certain kind of power. Note that this is not a line of argument I endorse – it is instead a sketch of the kind of argument that I think would be needed to establish plenitude.

We are in need of independent motivation for supposing that there are powers for propositions such as <2+2=4>. These powers must manifest necessarily, without stimuli or partner powers, and they cannot be powers to bring about events. I shall now suggest that such powers could be motivated by appealing to the notion that truthmaking is, in a sense, causal. Certain recent approaches to truthmaking treat it as a form of grounding. This avoids problems with purely modal approaches, according to which any entity that necessitates the truth of a proposition is a truthmaker for that proposition. On such approaches, anything is a truthmaker for a necessary proposition such as <2+2=4>. Intuitively, truthmakers should be relevant to what they make true – they should be among the entities that the proposition is *about*. Grounding-based approaches build in this kind of relevance in terms of the primitive notion that grounded entities, in addition to being necessitated by their grounds, exist *in virtue of* them.[17]

On a grounding-based approach to truthmaking, a true proposition <*p*> is true *in virtue of* the way the world is, and the part of the world that <*p*> is

[17] Further discussion is beyond the scope of this paper. See MacBride 2019 for full discussion.

about is the ground of <*p*>'s truth.[18] But now suppose we say, following Alastair Wilson, that grounding is *metaphysical causation* (Wilson 2018). This too is independently defensible. Grounds are sufficient for and metaphysically explain what they ground, and grounded entities depend on their grounds. Grounding and causation thus have several formal properties in common, and some have proposed analyses of grounding in terms of interventionist models of causation (Schaffer 2016; Wilson 2018). I need not commit to such claims for present purposes, since my aim is simply to point out that it is neither entirely implausible, nor without precedent, to treat grounding as a metaphysical form of causation. Now assuming truthmaker maximalism, <2+2=4> has a truthmaker.[19] If numbers are the truthmakers of <2+2=4>, and truthmaking is a form of grounding, and grounding is metaphysical causation, then the numbers are metaphysical *causes* of its truth. Now suppose further that causation of any kind always involves the manifestation of causal powers. It follows that numbers have a power for <2+2=4>, which they manifest in metaphysically causing its truth.

Assuming truthmaker necessitarianism – the view that truthmakers necessitate the truth of propositions they make true – truthmaking powers can't fail to manifest if instantiated. Given that the existence of the numbers is sufficient for the truth of arithmetic propositions, their truthmaking powers must be powers that manifest come what may. Metaphysical causal powers could also be said to have partners, in cases where the given truth refers to more than one entity. For truths like <2+2=4>, the most natural thing to say may be that the numbers 2 and 4 possess reciprocal powers to make the given proposition true, and manifest these powers mutually. The identification of grounding with metaphysical causation is controversial, and we are still in need of an argument that powers are involved in distinctively metaphysical causal relations. However, whatever its plausibility, the argument given above is an independent argument for plenitude. We don't need to commit to the claim that things have the power to have their necessary intrinsic properties, because we can argue directly that true propositions – both necessary and otherwise – are the manifestations of metaphysical causal powers to make true.

Instead of trying to defend, as Vetter does, the claim that numbers have noncausal powers *to have the features* in virtue of which they make <2+2=4> true, we should cut out the middleman and say instead that they have a meta-

[18] The first defence of this theory of which I am aware is in Rodriguez-Pereyra 2005. Several others have since defended variants of this claim. For more on the recent debate, see Trogdon forthcoming.

[19] Truthmaker maximalism is of course highly controversial; see Cameron 2008 for discussion of the main reason why sceptics tend to be sceptical – that is posits truthmakers for negative truths – and a defence of maximalism based on the claim that such truths are made true by the world, construed as having all its properties essentially.

physical causal power to make <2+2=4> true. There's nothing to stop us holding in addition that they have this metaphysical power in virtue of their necessary intrinsic features, but we don't need to say in addition that they have the power to have those features. The two proposals agree about the bearers of the relevant powers: on both accounts it is <*p*>'s truthmaker *x* that has a power for <*p*>. On Vetter's account, the power in question is the power to have the features in virtue of which *x* makes <*p*> true, while I am suggesting it is *x*'s power to make <*p*> true. Given truthmaker maximalism, the present proposal entails that for any true proposition <*p*>, something has a metaphysical power for <*p*>, because something makes <*p*> true. Whatever makes some proposition <*p*> true is the bearer of a metaphysical power for <*p*>, which it manifests in making <*p*> true, and which is just what, according to broad dispositionalism, is required to explain <*p*>'s possibility.

Let's compare this proposal to disjunctive dispositionalism, according to which <*p*> is possible just in case either <*p*> is true, or there is an efficient causal power to bring it about that *p*. On this theory, necessary truths are possible just because they are true, not because anything has a power for them, and powers take care of the possibility of contingently false propositions. On the face of it, broad dispositionalism is to be preferred to disjunctive dispositionalism due to its increased unity, but appearances can be deceptive. We can write the disjunctive dispositionalist possibility operator as follows:

DD: $\Diamond p \Leftrightarrow \{p \vee \exists \phi \rangle [p](\phi)\}$

Here '$\rangle[p]$' is shorthand for the predicate 'is an efficient causal power to bring it about that *p*', and '$\exists \phi \rangle [p](\phi)$' should be read as "something has, or some things have, an efficient causal power to bring it about that *p*",[20] The broad dispositionalist proposal outlined above appears more unified, since it grounds all possibility in powers, and permits a simple definition of the possibility operator according to which <*p*> is possible just in case something has, or some things have, a power for <*p*>. However, this obscures the fact that the theory posits two kinds of power: metaphysical and causal. Using the same notation as above for efficient causal powers, and using '$_{T}[p]$' for the predicate 'is a metaphysical power to make <*p*> true', we can write the broad dispositionalist definition of the possibility operator as a disjunction as well:

BD: $\Diamond p \Leftrightarrow \{\exists \varphi_{T}[p](\varphi) \vee \exists \phi \rangle [p](\phi)\}$

[20] Notation from Yates 2015. There I do not explicitly restrict quantification to efficient causal powers, but I did tacitly assume that all powers were efficient causal, so here I make the restriction explicit in the definition.

The question now is: what we really gain by moving from (DD) to (BD)? We had to make several controversial assumptions to argue for metaphysical causal powers, so (BD) had better have significantly greater theoretical unity than (DD). However, the power to make a given proposition true is a very strange kind of power, and doesn't seem to resemble efficient causal powers at all, at least as they are typically conceived by proponents of powers ontologies. As we saw in §1, efficient causal powers are typically said to be individuated by their places in a second-order causal structure. Powers to make true, by contrast, don't seem to involve much individuative structure at all – every such power is directed at the propositions it makes true, and necessitates the truth of those propositions just by being instantiated. According to (DD), possibility is grounded either in truth, or in the power to bring about truth; while given (BD), possibility is grounded either in the power to make true, or in the power to bring about truth. On reflection, it's not obvious that (BD) is more unified than (DD), due to the fact that (BD) grounds possibility in two very different kinds of power. The unity of a powers theory of modality depends not only on whether the properties to which we appeal as the grounds of modal truths are all called "powers"; it also depends on the unity of the set of powers to which we appeal. If metaphysical powers are powers in name only, then broad dispositionalism appealing to both efficient causal and metaphysical powers is *unified* in name only.

Against this it might be said that there is independent reason for thinking that metaphysical powers are similar enough to efficient causal powers for (BD) to constitute a unified theory. If grounding is metaphysical causation, and truthmaking is a form of grounding, then it's no major leap to hold that whatever makes <2+2=4> true has a genuine causal power to do so, even though it can't be the power to cause an event. Such powers manifest necessarily, and don't have stimuli or partner powers, but as Vetter points out, there are efficient causal powers with those features. Much turns on how persuasive the case for metaphysical powers is, and given the commitments we needed to argue for them, the certainly aren't cheap: (i) truthmaking is a form of grounding, (ii) grounding is metaphysical causation, (iii) all causation involves the manifestation of powers, (iv) truthmaker maximalism. The happier one is to embrace these commitments, I suspect, the more likely one will be to see (BD) as significantly more unified than (DD). Conversely, dispositionalists who find the idea that truthmaking is causal dubious, or who accept that claim but have reservations about the idea that this kind of causation involves the manifestation of powers, will likely see a world of difference between the efficient causal power of a brick to break the window, and the metaphysical power of the shattered glass to make <the window is broken> true.

4. *Conclusion*

Dispositionalism is the claim that modality is grounded in the powers of things. Narrow dispositionalism is the combination of dispositionalism with the claim that all powers are efficient causal, but it suffers from serious problems of material and formal adequacy, since there are no efficient causal powers available to ground the possibility of propositions like <2+2=4>. My attempt to solve this problem in Yates (2015) was to embrace disjunctive dispositionalism, according to which the possibility of contingently false propositions is grounded in powers, with the possibility of truths following trivially from their truth. Vetter's solution is to embrace broad dispositionalism, according to which there is a plenitude of both efficient causal and noncausal powers available to ground possibilities.

Against the charge that plenitude is ad hoc, Vetter argues that dispositions come in degrees, with the necessary intrinsic features of objects being limiting cases of dispositionality. Powers for propositions such as <2+2=4> are then grounded in the powers of their truthmakers to have the necessary intrinsic features in virtue of which they make those propositions true. However, Vetter's argument from degrees fails to establish plenitude. Although it's plausible that dispositions come in degrees, there are clear cases of disposition pairs – such as the disposition to sing when it rains and the disposition to not sing when it rains – such that it's possible to instantiate neither of them to any degree. Hence, it doesn't follow from *x*'s having P intrinsically and necessarily that *x* has a power for P. As an alternative way of defending plenitude, I have suggested in the present work that if truthmaking is a kind of metaphysical causation, then we can argue for the existence of metaphysical causal powers to ground the possibility of truths that would otherwise seem beyond the reach of causality. Necessary truths are beyond the reach of efficient causation, but given truthmaker maximalism, no truth is beyond the reach of metaphysical causation, so for every truth, there is a power.[21]

David Yates
david.yates@campus.ul.pt
FCT Researcher of LanCog, Centre of Philosophy, University of Lisbon

[21] I am particularly grateful to Barbara Vetter for our earlier discussions of her argument for plenitude, and to three anonymous referees for some very helpful comments and criticisms. I gratefully acknowledge the financial support of the fundação para a ciência e a tecnologia (award #s PTDC/FER-HFC/30665/2017 and IF/01736/2014).

References

Bigaj, T., 2017, "On Quantum Entanglement, Counterfactuals, Causality, and Dispositions", in *Synthese*, https://doi.org/10.1007/s11229-016-1304-9.

Bird, A. 2007a, *Nature's Metaphysics*, Oxford University Press, Oxford.

Bird, A. ,2007b, "The Regress of Pure Powers", in *Philosophical Quarterly*, 57: 513-534.

Cameron, R., 2008, "How to be a Truthmaker Maximalist", in *Noûs*, 42: 410-421.

Contessa, G., 2010, "Modal Truthmakers and Two Varieties of Actualism", in *Synthese*, 174: 341-353.

Ismael, J., and Schaffer, J., 2016, "Quantum Holism: Nonseparability as Common Ground", in *Synthese*, https://doi.org/10.1007/s11229-016-1201-2.

Kim, J., 1976, "Events as Property Exemplifications", in M. Brand and D. Walton (eds.), *Action Theory*, D. Reidel, Dordrecht: 310-326.

Jacobs, J.D., 2010, "A Powers Theory of Modality – Or, How I Learned to Stop Worrying and Reject Possible Worlds", in *Philosophical Studies*, 151: 227-248.

MacBride, F., 2019, "Truthmakers", in *The Stanford Encyclopedia of Philosophy*, E.N. Zalta (ed)., https://plato.stanford.edu/archives/spr2019/entries/truthmakers.

Martin, C.B., 2008, *The Mind in Nature*, Oxford University Press, Oxford.

Nolan, D., 2015, "Noncausal Dispositions", in *Noûs*, 49: 425-439.

Pruss, A.R., 2011, *Actuality, Possibility and Worlds*, Continuum, New York.

Rodriguez-Pereyra, G., 2005, "Why Truthmaking?", in H. Beebee and J. Dodd (eds.), *Truthmakers: the Contemporary Debate*, Oxford University Press, Oxford: 17-31.

Schaffer, J., 2016, "Grounding in the Image of Causation", in *Philosophical Studies*, 173: 49-100.

Trogdon, K., forthcoming, "Grounding and Truthmaking", in *The Routledge Handbook of Metametaphysics*.

Vetter, B., 2014, "Dispositions Without Conditionals", in *Mind*, 123: 129-156.

Vetter, B., 2015, *Potentiality: From Dispositions to Modality*, Oxford University Press.

Vetter, B., 2018, "A plenitude of powers", in *Synthese*, DOI, 10.1007/s11229-018-1758-z

Williams, N.E., and Borghini, A., 2008, "A Dispositional Theory of Possibility", in *Dialectica*, 62: 21-41.

Wilson, A., 2018, "Metaphysical Causation", in *Noûs*, 52: 723-751.

Yates, D., 2015, "Dispositionalism and the Modal Operators", in *Philosophy and Phenomenological Research*, 91: 411-424.

Yates, D., 2018, "Inverse Functionalism and the Individuation of Powers", in *Synthese*, 195: 4525-4550.

Potentiality: actualism minus naturalism equals platonism

Giacomo Giannini, Matthew Tugby

Abstract: Vetter (2015) develops a *localised* theory of modality, based on potentialities of actual objects. Two factors play a key role in its appeal: its commitment to Hardcore Actualism, and to Naturalism. Vetter's commitment to Naturalism is in part manifested in her adoption of Aristotelian universals. In this paper, we argue that a puzzle concerning the identity of unmanifested potentialities cannot be solved with an Aristotelian conception of properties. After introducing the puzzle, we examine Vetter's attempt at amending the Aristotelian conception in a way that avoids the puzzle, and conclude that this amended version is no longer to be considered naturalistic. Potentiality theory cannot be both actualist and naturalist. We then argue that, if naturalism is to be abandoned by the actualist, there are good reasons to adopt a Platonist conception of universals, for they offer a number of theoretical advantages and allow us to avoid some of the problems facing Vetter's theory.

Keywords: Platonic universals; potentiality; naturalism; actualism; dispositionalism.

1. *Introduction: Vetter's potentiality-based theory of modality*

Vetter's (2015) potentiality theory of modality is the view that the truth and falsity of alethic modal claims – in particular claims about metaphysical modality – have their source in the potentialities present in the actual world. Potentialities are the (only) truthmakers of modal discourse. According to the view, the potentialities of actual concrete objects (past and present) fix all the modal facts and determine the topology and extension of the modal space.

For the purposes of this paper, we can formulate the potentiality theory as the conjunction of the following two theses:

PPoss: 'possibly *p*' is true iff and because there is some potentiality whose manifestation, if manifested, would make '*p*' true.

philinq VIII, 1-2020, pp. 117-140
ISSN (print) 2281-8618-ETS DOI: 10.4454/philinq.v8i1.278

PNec: 'necessarily *p*' is true iff and because there is no potentiality whose manifestation, if manifested, would make 'not-*p*' true.[1]

This is a bare-boned version of Vetter's theory, one which could hardly hope to achieve extensional correctness, that is, to generate enough modal truths. Vetter ingenuously enhances her account in a number of ways to remedy this, introducing joint potentialities (neither Fischer-Dieskau nor Gerald Moore could perform Schubert's *Die Winterreise* on their own, but could do so together) and extrinsic potentialities (Fischer-Dieskau has the potentiality to perform *Die Winterreise* with Gerald Moore). Building on Borghini and Williams' concept of a branching disposition (2008: 32), Vetter also develops the important notion of iterated potentialities (I do not have the potentiality to speak Finnish, but I have the potentiality to learn how to speak Finnish – inelegantly, I have the potentiality to have the potentiality to speak Finnish). In this paper we will only discuss the minimal version of the theory, but it is important to keep in mind that it is because of the development of these further aspects (and many others, such as the treatment of gradability and the development of a rigorous semantics) that Vetter's work truly stands out as the canonical text on dispositions for many years to come.

There are three factors that crucially contribute to the appeal of Vetter's potentiality theory as an overall account of modality, which can be summarised under the headings of Realism, Hardcore Actualism, and Naturalism.

Vetter (2015: 33-60) presents original and persuasive arguments against the reduction of dispositions to counterfactual conditionals, based on the gradability of dispositions, which reinforce and round out the classic, well-known[2] objections based on finks and antidotes: irreducibly modal properties are part of the furniture of this world. We have to be realist about them. A tempting thought ensues: since we have to accept these properties into our ontology, we might as well make them do as much work as possible – maybe we could even explain *the whole of modal discourse* with them. However, the potentiality theory does not aim to be a reductive theory of modality, in the way that, say, Lewis's (1986) account is. Rather, it aims at rearranging the landscape: according to the potentiality theorist, we should take *localised* modality[3] to be more fundamental than the

[1] These can be formalised by expressing potentiality with a predicate modifier, POT: '$\Diamond p$' is true iff POT[Φ](xx), where Φxx would make '*p*' true.

[2] See Martin (1994) and Bird (1998), as well as Manley and Wasserman (2008).

[3] "A potentiality is localised in the sense that that it is a property of a particular object […] possibility, on the contrary, is not localised this way. Its being possible that such-and-such is not primarily a fact about any one particular object; it is a fact about how things in general might have turned out to be" (Vetter 2015: 2). In linguistic terms, the difference can be expressed by the fact that "the argument places [of non-localised modal operators] must always be filled by an entire sentence […] the

non-localised one, just like essence is prior and more fundamental than necessity according to Fine (1994); attempts at reducing dispositions to counterfactuals fail in part because they get the order of the explanation wrong. Indeed, the Potentiality Theory and Essentialism are very similar – the main difference being that the former is a "possibility-first" and the latter a "necessity-first" theory of localised modality.[4] The hope is that the unavoidable realism about dispositions and potentialities might lead to ontological parsimony elsewhere.

And indeed, one of the main features of Potentiality Theory (as well as Essentialism) is that both aim to be what Contessa (2008) has dubbed "Hardcore Actualist" theories, that is, to reject the idea that possible worlds (however conceived) should play a role in making modal statements true – in short, they reject the idea that the Leibnizian biconditionals are *metaphysically* informative, as it were. One of the purported advantages of this approach is that we need not be committed to strange and controversial entities such as possible worlds, since they play no role in fixing the modal truths. This is often taken to be a gain both in ontological parsimony and common sense; the Potentiality theorist offers the prospect of making sense of modality with a lightweight, safe and sane ontology: all we need are powerful actual objects. If we are to take the label 'Actualism' seriously, we should then think that Potentiality Theory is committed to the "Being Constraint" (Williamson (2013: 148)[5] which is best expressed by the generalisation (that is, including higher-order variables) of the following two theses:

1. $\Box\forall x\,\Box(Fx \rightarrow \exists z\, x{=}z)$
2. $\Box\forall x\Box\forall y\,\Box(Rxy \rightarrow (\exists z\, x{=}z \wedge \exists z\, y{=}z)$

In English, (1) says that necessarily, for all x, necessarily if x has a property F, then x is something. (2) says that necessarily, for all x, and necessarily for all y, necessarily if x stands in a relation R to y, then x is something and y is something.

The third noteworthy feature of Potentiality Theory is that it promises a naturalistic account of modality. This point is somewhat implicit in the com-

operators for localised modalities, on the other hand, must have at least one argument for the *object* (or objects) to which the modality belongs, and another argument place for that which is intuitively the content of the modality, and which is most naturally expressed by a predicate" (Vetter 2015: 5).

[4] These terms are usually used to refer to different approaches to the epistemology of modality; here we employ them in a metaphysical sense: for the potentiality theorist, the fundamental phenomena are what make possibility statements true, and necessity is to be obtained from there. It is tempting, but not necessary, to think that the epistemology would be isomorphic.

[5] As Williamson notes, the Being Constraint can be seen to capture what Plantinga (1983: 11) meant by "Serious Actualism".

mitment to Hardcore Actualism, and often the two are run together, but we think they contribute to the development of Vetter's theory in different respects and can pull it in different directions, and so are best kept apart. In expounding the appeal of a localised theory of modality, Vetter mentions that such an account promises to provide an account of possibility and necessity that anchors them to

> just the ordinary objects of this, the actual, world, with which we are in regular epistemic contact […] if it succeeds then it does so by anchoring possibilities in realistically respectable bits of the world, ordinary concrete objects. (Vetter 2015: 11)

The "respectability" of Potentiality Theory can be attributed to two main factors: on the one hand, its vindication of a familiar and commonsensical Aristotelian ontology of objects and properties, and on the other hand, the fact that such objects are epistemically accessible: "actual objects, with which we have epistemic contact" (Vetter 2015: 11). Note: it is not only the fact that such objects are *actual* that makes them "respectable" and accessible: it is also the fact that they are *concrete* – spatiotemporally, and hence causally, linked to us. In order to learn about modality we can use a "powerful telescope" (Kripke 1980: 44), after all! We think that this is sufficient evidence to attribute to Potentiality Theory a commitment to Ontological Naturalism, which can be characterised as "the doctrine that reality consists of nothing but a single all-embracing spatio-temporal system" (Armstrong 1981: 149).

2. *Unmanifested potentiality and the fundamental puzzle*

As we have seen, Vetter's theory of modality aims to be both hardcore actualist and naturalistic: modality is ultimately a matter of how concrete objects are. Whilst we agree that a strongly naturalistic actualism is desirable, we are (with regret) doubtful that naturalism can provide enough ontological resources for a coherent metaphysical account of irreducible potentialities. Our aim in this section is to articulate those doubts and to conclude that a Platonic approach to potentialities (or something like it) provides the ontological resources that we need. As far as we can tell, a move to Platonism about properties would accommodate many of the details of Vetter's potentiality-based account, and it can still be considered as an actualist theory. Moreover, as we shall see in the next section, Platonism allows us to ground more possibilities than Vetter can allow, and therefore overcomes some of the objections facing her theory and displays superior theoretical virtues.

Vetter's potentiality theory of modality is to be considered naturalistic, in part, because it goes hand-in-hand with a broadly Aristotelian approach to

properties. On this view, properties (or 'universals') do not exist independently of their concrete instantiations – of how concrete things are. Unlike the Platonists, Aristotelians do not have to say that properties exist outside of space and time. Rather, Aristotelian universals exist entirely through the concrete things that instantiate them, *in rebus.* The dependence of Aristotelian universals on their instantiations is typically captured by what Armstrong calls the 'Principle of Instantiation', which for each property

> Demand[s] that it is a property of some particular[, and f]or each relation universal [it must] be the case that there are particulars between which the relation holds. (Armstrong 1989: 75, quoted in Vetter 2015: 271)

The ontological dependence of Aristotelian universals on their instantiations is fairly weak in the sense that it is generic: in order to exist, a property must be a property of *some* particular, but it doesn't matter which particular performs the job. Nonetheless, the dependence of properties on their concrete instantiations is strong enough to rule out the existence of type-uninstantiated properties, given that such properties are not properties of anything concrete whatsoever.

The problem presented in this section can be summarised as follows: if we do not have uninstantiated properties in our armoury, then it is difficult to make metaphysical sense of potentialities whose manifestations are never manifested or actualised. We will first illustrate the difficulty, and show how forever unmanifested properties are problematic for the Aristotelian conception of universals. We will then show that Vetter's amended account does not succeed in solving the problem *while* retaining its commitment to naturalism, and is therefore in no better position than Platonism in this regard. We finally argue that Platonism offers a number of theoretical advantages, and should therefore be preferred to an immanentist conception of properties.

2.1. Directedness Platitude

Versions of problem we now turn to have been discussed previously (Armstrong 1997; Molnar 2003; Bird 2006; Tugby 2013) and are acknowledged by Vetter herself (2015: Sect. 7.5). In our view, the problem is best described as a tension between the following two platitudes.[6] On the one hand, potentialities are (partially) individuated by their manifestation types: the identity of a po-

[6] We adopt the term 'platitude' in order to be terminologically consistent with Tugby (2013). By it, we mean that both 'platitudes' express what we take to be central, and in our view non-negotiable, features of powers metaphysics. Of course, defenders of certain versions of the Powerful Qualities view could reject DP, and therefore deem the term 'platitude' to be unwarranted. Thanks to an anonymous referee for highlighting this.

tentiality is determined by what it is a potentiality *for*. In order to explain *what mass is*, one can do nothing more than to say that massy objects are disposed to exert gravitational force on one another.[7] And to explain *what something is*, is precisely to show what *individuates* it. This point is often expressed by saying that dispositions are essentially "directed" towards certain manifestations rather than others. Tugby (2012: 168) dubs this the 'Directedness Platitude':

DP: Dispositions are directed towards their manifestation properties and it is in virtue of this directedness that the identity of a disposition is fixed.

To be clear, we are talking about individuation in the metaphysical sense rather than, say, the cognitive sense. Cognitive individuation is a mind-dependent act that involves singling out an entity in thought. Although it is true that we cognitively individuate dispositions by thinking about the manifestations that they are dispositions for, this is not the kind of individuation that is at issue above. Even if minded creatures were not to exist, it would still be the case that the nature of a disposition is determined by the type of manifestation that it is a disposition for. Thus, in speaking of the individuation of dispositions, we are speaking of a mind-independent *metaphysical* determination relation that distinguishes a given dispositional property from all other possible dispositional properties. As Lowe puts it, metaphysical individuation is "the relation that obtains between entities *x* and *y* when *x* determines or 'fixes' (or at least *helps* to determine or 'fix') *which* entity of its kind *y* is" (2010: 9).

By stating that the identity of powers is *partially* determined by what they are *for*, we do not wish to be committed to the idea that the nature of a dispositional property has to be *exhausted* by what it is for, or its manifestation relation (as in Mumford 2004). Other factors, such as its stimulus conditions (Bird 2007) or its reciprocal disposition partners, or even its degree and granularity (Vetter 2015), may play a prominent role.

It is less clear whether proponents of a 'Powerful Qualities' view of powers can subscribe to the Directedness Platitude. The core idea of such a view is that dispositional properties are *both* powers and qualities; they are *both* categorical and dispositional. There are two crucial issues to be settled before we can understand whether DP also applies to Powerful Qualities:

1. How to characterise what are qualities, or categorical properties
2. How to characterise the relation between the powerful and the qualitative aspect

[7] We assume here that mass is a *fundamental* property which cannot be further reduced nor analysed.

The crucial point in assessing whether DP applies also to Powerful Qualities views concerns how qualities are characterised. For instance, according to Bird (2016) the difference between a power and a categorical property lies in the fact that the former has its modal profile fixed, whereas the latter is 'modally variable'. This is due to the fact that "a power is an ontic property a) that has a dispositional essence", and b) "[…] whose identity is given by its causal/dispositional/nomic role" (Bird 2016: 347). Presumably this would mean that a categorical property *does not* have its identity given by its causal/dispositional/nomic role. That is, DP applies to powers and not to categorical properties. If the Powerful Qualities view states that powers are *also* categorical, this would mean that Powerful Qualities *also* have a primitive identity: presumably, then, DP would not apply to them.

But if we accept Bird's characterisation of the difference, and if we take the relation between powerful and qualitative aspects to be identity (Heil 2003; arguably, Martin 2009) then it would follow that Powerful Qualities views are trivially false: the very same property would have both a fixed and variable modal profile, and its identity would be both primitive and fixed by its nomic role. Obviously, this cannot be the characterisation of the distinction adopted by the proponents of the view.

Taylor (2013: 93-94), on the other hand, elucidates the difference thus: "a dispositional property is any property to which it is essential that it conveys upon the object that instantiates it the power to behave in a certain way given certain stimuli", whereas "qualitative/categorical properties essentially contribute to the overall makeup of how an object is now". This characterisation does not impinge on differences on identity conditions, and obviously does not immediately lead to contradiction when paired with Identity Theory. Nothing in the characterisation of categorical properties entails that their identities are primitive, and that DP should not apply to them.

Similarly, Heil (2012: 59) seems to suggest that the defining feature of categorical properties is just that they are actual or occurrent: "qualities are here and now, actual, not merely potential, features of the objects of which they are qualities", and Strawson even equates 'categorical' with 'being': "all being is categorical being because that's what it is to be! That's what being is!" (Strawson 2008: 278). Again, this understanding of categorical is neutral about the identity conditions of qualities, and so does not entail that DP does not apply.

Obviously, there are many versions of the Powerful Qualities view, articulated *inter alios* by Martin (1997; 2009), Heil (2003), Jacobs (2011), Ingthorsson (2013; 2015), Taylor (2013; 2018) Giannotti (forthcoming), and Contessa (2019). These authors often differ over the details of the answers offered to i) and ii), and so a satisfying answer could be given only after a detailed examination.

Such a project unfortunately goes beyond the scope of this paper. Perhaps not all of these versions are compatible with DP. If it is shown that a Powerful Quality theory which relies on the elucidation of the categorical/dispositional property distinction as involving a radically different theory of individuation is viable and preferable to the other Powerful qualities views,[8] then this will be a serious competitor to our theory below, and more work will have to be devoted to establish which theory has the upper hand. For the time being, we have shown that the Directedness Platitude is compatible with at least some versions of the Powerful Qualities view. So, our argument concerns at least a considerable portion of the theories of powers on the market, including Vetter's own theory which we discuss below.

2.2. Central Platitude

On the other hand, in Tugby's 2013, it is argued that a *desideratum* for any metaphysics of potentiality (or "dispositionalism"[9]) is that it accommodates what Molar (2003) referred to as "Independence", and what Tugby (2013) called the "Central Platitude":

CP: A particular can have a disposition even if it never manifests that disposition (2013: 454).

We take it that this platitude is part and parcel of any realist approach to dispositions. To say that a disposition is a real property of things is precisely to say that it can be instantiated even if it is not being exercised. Although the *manifestation* of a disposition is dependent on certain situations obtaining, such as a soluble object being placed in water, the disposition itself (e.g. water-solubility) is not itself dependent on such situations arising. If we are realists about dispositions, then we should surely accept that a piece of salt is water-soluble even if the salt (or any other substance) never finds itself in water. Even if salt's water-solubility is never actually manifested, the disposition still makes a counterfactual difference to the world: the disposition is ascribable precisely because salt *can* dissolve even if it doesn't actually dissolve.

Here is the tension between the platitudes: Per **DP** and the Being Constraint, if a potentiality is directed to its manifestation, then the manifestation

[8] Possibly on the basis that it would offer a solution to the regress of pure powers discussed *inter alios* by Lowe (2010), Bird (2007), Ingthorsson (2015). However, this might not be so easy: Taylor (2018) has argued that an Identity Theory of Powerful Qualities is not better off than a pure powers view in this respect.

[9] In what follows we shall use the terms 'dispositions' and 'potentiality' interchangeably. As mentioned earlier, Vetter's notion of potentiality is broader than the notion of a disposition, but this difference is not important for the purposes of the following arguments.

exists. But, per **CP**, the manifestation needs not occur. Armstrong (1997: 78) concluded that dispositions either give raise to a contradiction, or are committed to a "Meinongian" ontology, where dispositions are somehow related to the non-existent. Given the commitment to Hardcore Actualism, both results would be equally unwelcome. Fortunately for the friends of potentialities, Armstrong's argument is too quick and is based on an ambiguous reading of 'manifestation'. 'Manifestation' can stand both for the second *relatum* of directedness (what a potentiality is *for*) and for its obtaining – the fact that the potentiality has been successfully exercised, and brought the concrete manifestation about. A contradiction arises only if we read 'manifestation' in both **CP** and **DP** as referring to the existence of what the power is *for*, thus reading **CP** as stating that the object the disposition is for can fail to *exist*.

But there is no reason to read **CP** as concerning the *existence of the manifestation, understood as what a disposition is for.* What **CP** maintains is that such manifestation can fail to occur, or to be brought about: the manifestation (object) can fail to be manifested (be brought about). This suggests the following schema to escape the problem:

1. $\Diamond F(P) \wedge \neg F(M)$	(Central Platitude)
2. *Directed* (P, M)	(Directedness Platitude)
3. $\exists X\ (X{=}M)$	(2., Being Constraint)
4. $\Diamond\exists X\ (X{=}M \wedge \neg F(\mathrm{M}))$	(1,3)

The platitudes generate a contradiction only if *F* is taken to be 'existence', and results in possibilism (i.e., the existence of non-actual entities) only if *F* is taken to be 'actual'.[10] But there is no reason to do so: there is a variety of other relevant properties that *F* could be.[11] In particular, it was suggested by Mumford (2004: Ch. 11) that *F* could be 'instantiation', and that the manifestation could be a universal: in this case **CP** would just state that the disposition could fail to bring about the instantiation of the universal on some particular occasion, while its identity is fixed by the universal, as per **DP**. This solution preserves Hardcore Actualism (for it accepts the Being Constraint and admits only actual objects, such as universals) and, in so far as the universals in question are Aristotelian (spatiotemporally located), Naturalism. We take it that

[10] See Bird (2006, 2007).

[11] For instance, one could in principle characterise *F* to be 'concreteness' (or something like that), claiming that unmanifested manifestations display a kind of "purely logical existence" (Williamson 2002); one of us favours a solution along those lines. For the purposes of this paper, however, we will assume that 'instantiation' is the best way to articulate the proposal, and only consider universals as manifestations.

Vetter (2015: Ch. 7) adopts universals as manifestations for potentialities precisely in order to avail herself of Mumford's strategy to solve the tension and maintain a commitment to both Hardcore Actualism and Naturalism while retaining **DP**.

The problem for this approach is that the Central Platitude also establishes that all instances of a potentiality property could go unmanifested: it is possible that the Aristotelian manifestation universal is never instantiated, and therefore does not exist. This can easily be shown: either there is something special about the unmanifesting instances, or there is not. If there is not, then one has to admit that *any* instance (token) of non-maximal potentiality could fail to bring about its manifestation. But if this were the case, surely it could be the case that *all* instances of potentialities happen to fail to bring about their manifestations: since the unmanifesting instance was arbitrary, we can't block the universal generalisation: if *F* is true of an arbitrary *x* of the domain, then $\forall x Fx$. So, whole potentiality types can fail to manifest.

The other option is to say that there is something special about certain instances. Let's call this distinguishing feature '*K*', and maintain that only a subset of the token potentialities happen to have it. Then it would seem that Token-Independence (to adopt the terminology of Molnar 2003) is not a feature of *potentialities per se*, but rather of *K*: *K* is the difference-maker and hence **CP** applies only to *K*-potentialities. But this clashes with the Central Platitude, which was supposed to be a principle of dispositions *qua* dispositions.

Vetter (2015: 85-94) allows potentialities of maximal degree that necessarily bring about their manifestations, and hence will reject the idea that the Central Platitude obtains in virtue of the nature of potentialities *qua* potentialities – rather, she will maintain that **CP** holds in virtue of the degree of a potentiality. Still, the problem remains, although on a slightly more limited scale: even if we think that **CP** does not hold for all potentialities, it will hold for all the non-maximal potentialities, which can be Type-Independent.

Call the universals that are never manifested 'aliens', and potentialities that are directed to them 'alien potentialities'. Given that Aristotelian universals ontologically depend on their instantiation, if they are never instantiated they do not exist. This leaves us unable to adopt Mumford's solution in the case of aliens. We are then left with what we call the alien puzzle of potentiality: how are alien potentialities individuated? It is difficult to see how this puzzle can be solved without renouncing Hardcore Actualism, or renouncing Naturalism.

One could salvage Naturalism and Aristotelian universals by denying the validity of the derivation from 2. to 3. This is the move that the defenders of "physical intentionality" make (e.g. Molnar 2003). But that would mean violat-

ing the Being Constraint and hence force Vetter to abandon her commitment to Hardcore Actualism. We take it that Vetter would find this solution to be highly undesirable, and we wholeheartedly agree.[12]

Alternatively, one could renounce (ontological) Naturalism and adopt Platonic Universals: since these do not depend upon their instantiations for their existence, it is perfectly acceptable to think that there are uninstantiated (alien) universals – no modification of the account would be required to make sense of alien potentialities. But Vetter seems to think that this would be an unacceptable loss: Naturalism plays an important part in making the ontology of Potentiality Theory "respectable", because

> the picture of metaphysical modality that I am offering derives its attraction, I believe, not merely from the fact that it thereby locates modality in properties, viz. potentialities, but also from the fact that it thereby locates modality in *objects* of the ordinary kind: concrete objects. (Vetter 2015: 270)

So Vetter finds herself in the uneasy situation of accepting the force of the fundamental puzzle for the Aristotelian potentiality theorist (2015: 271), without being willing to give up either Actualism or Naturalism. In order to reconcile Actualism with Naturalism, Vetter proposes to modify the Principle of Instantiation discussed earlier in a way that can ground talk of unmanifested properties within an Aristotelian framework. The modified principle, called the 'Principle of Potential Instantiation' (PPI) is as follows:

> **PPI:** Every universal must be at least potentially instantiated: there is a property universal of being F only if there is some particular thing which is F, is potentially F, or is potentially such that something is F; there is a relation universal of R-ing only if there are some particular things which R, or which potentially R, or which are potentially such that some things R. (Vetter 2015: 272)

Vetter then explains how this modified principle of instantiation gets around the problem of unmanifested dispositions:

> … the claim that something has a potentiality to have (or produce or constitute something which has) the actually uninstantiated property of being F is not in jeop-

[12] Note that it is not wholly clear that such a move would preserve Naturalism, either: it is far from clear what it means for a physical state to be directed towards something that does not exist. In the mental case, we can at least say that the non-existent intentional objects are represented and that those representations serve to individuate different mental states. But since Molnar (2003) maintains that physical intentional states do not represent, it is not obvious how such intentional states are individuated; the risk is that physical intentionality *de facto* makes the individuation of potentialities a primitive, undetectable fact. But this would violate DP, as well as make them epistemically inaccessible.

ardy because there might be no property of being F. Rather, that claim, if true, guarantees that there is such a property, because this is precisely what it takes for there to be a property of being F. (2015: 272).

This is a striking quotation because on a natural reading of this proposed solution, ontological naturalism seems to be under threat. Both **PPI** and the application of it employs existential quantification: it is said *there is* a property universal only if there is some particular thing which is *F*, is potentially *F*, or is potentially such that something is *F*. Crucially, in cases where something is potentially *F*, but nothing ever manifests that potentiality, then *F* surely remains *uninstantiated*. Hence, on a natural reading of this principle, it commits us to the existence of *uninstantiated properties.* The approach remains Aristotelian, in so far as it maintains that universals depend on concrete entities for their existence: not only their instances, but also the potentialities that are directed at them. The problem is that, it seems to us, such Aristotelianism is no longer naturalistic. It was not the fact that properties are dependent upon something that made them naturalistically *kosher*, but rather the fact that they were located in space-time: they were located where their instances were. But now consider the uninstantiated Aristotelian universals allowed by **PPI**: where in space-time are they? Surely they are not located where their instances are, because there are none. But they cannot be located where the potentialities directed to them are, either, because i) nothing grants that the manifestation is co-located with the potentiality, and ii) if they were located there, then by Aristotelian lights they would surely be *instantiated* by the bearers of the potentiality, which would catastrophically mean that every power is always already manifested.

If the Aristotelian universals allowed by **PPI** are not located in space-time, then a theory invoking them is not compatible with Naturalism as we understand it. But if this is the case, it is no longer clear how Vetter's approach differs from Platonism. We take the allowance of properties which are not instantiated to be the core commitment of Platonism and one that is inconsistent with the ontologically naturalistic commitment to locating all entities in space and time. Vetter's position seems dangerously (or, we think, fortunately) close to Platonism.

Later on Vetter distances herself from Platonism by associating Platonism with the view that there are super-alien properties, where super-alien properties are properties that "no actual thing ever had a potentiality to have, to produce, or to constitute" (2015: 69). However, we do not see why all Platonists have to accept the existence of super-alien properties. As we shall see in section 3, we think there are good reasons for accepting that there are super-alien properties. Nonetheless, a version of Platonism which follows Armstrong's "a posteriori realism" about universals (1978), and uses science as guide as to what

(non-super) alien properties there are, is perfectly coherent. Again, we take the core commitment of Platonism to be that properties do not need to be instantiated in order to exist. Whether or not there are super-alien properties is a matter for in-house dispute between Platonists.

We can think of a couple of interpretations of the Principle of Potential Instantiation that avoid Platonism, as we understand it, but neither of them is appealing. First, it could be said that the existential quantifier employed in the **PPI** does not entail an ontological commitment to uninstantiated properties. This move would require us to distinguish between so called "quantifier commitment" and "ontological commitment" (see e.g. Azzouni 2004). However, the problem with this approach is that it is precisely the sort of move that is employed by some Meinongians, as a way of denying an ontological commitment to non-existent objects (see e.g. Priest 2005). Hence, this approach to uninstantiated manifestations would arguably place Vetter in the category of Meinongian dispositionalists. We agree with Armstrong that this is best avoided.

However, there is another option. Recall that the difference between Vetter's position and Platonism lies in the fact that Aristotelian universals are not ontological independent, whereas Platonic universals are. So, perhaps a more promising way for Vetter to avoid ontological commitment to uninstantiated properties is to say that unmanifested properties exist in the sense that they are *grounded* in the potentialities of things. This option is open to Vetter because she independently accepts that grounding is an important relation (2015: 26-28).[13] Grounding is a metaphysical determination relation that provides an explanation for why something is so (see e.g. Schaffer 2009 and Rosen 2010). Importantly, grounded entities can be thought to be derivative and "ontological free lunches", given that their being is fully explained by their grounds. One could think that the only ontological commitment that *really* matters concerns only *fundamentalia*. Hence, if uninstantiated manifestations were grounded in the relevant potentialities, quantifying over them would involve no increase in *genuine* ontological commitment. We would not have to say that uninstantiated manifestations exist over and above the instantiated potentialities, and therefore a full-blown Platonic commitment to uninstantiated manifestations would be avoided. We suspect that a view along these lines may be what Vetter ultimately has in mind.

[13] Note that it is far from obvious that Potentiality Theory is in the end compatible with grounding. Grounding is typically thought to be a modally-laded notion. For instance, the orthodox (e.g. deRosset 2010; Fine 2015) view of grounding is that if Γ (fully) grounds B, then it is metaphysically necessary that B exists if Γ does. It is not clear that Potentiality Theory can explain such modal consequences in a plausible way. For the sake of argument, we will not press this point and assume that the use of grounding is compatible with Potentiality Theory.

Unfortunately we think that this line of defence faces some thorny problems. The most serious difficultly concerns the original difficulty of individuating potentialities that are unmanifested. As we have seen, potentialities are plausibly individuated by their manifestation properties. If those manifestation properties are ontological free lunches, then we have a problem: for it is difficult to see how something that is not an ontological free lunch can be individuated by something that is itself an ontological free lunch (see e.g. Barker 2009: 247 and Tugby 2016: sect. 3.1 for related points). It seems incoherent to suppose that potentialities, which are ontologically fundamental, could be individuated by something less fundamental than themselves and which they themselves ground. To think otherwise is to violate the metaphysical analogue of what Sider calls the 'principle of purity' (2011: 106), which says that "fundamental truths involve only fundamental notions" (see also Jaag 2014: sect. 3 for a closely related discussion of purity in the context of dispositionalism).

To sum up this section, we have argued that Vetter's attempt to solve the fundamental puzzle of potentialities either fails, or she ends up precisely with a sort of ontologically non-naturalistic approach to properties that is close to Platonism. The choice between Aristotelian and Platonic universals is, of course, a metaphysical matter, and should be settled in the canonical way: weighing costs and benefits of the overall theory. The difference between a Platonist position and Vetter's **PPI** boils down to this:

> Platonism: Potentialities depend upon their manifestations (universals) for their identity. These universals can exist uninstantiated and do not depend upon concrete potentialities for their existence.
>
> Vetter's Aristotelianism: Potentialities depend upon their manifestations (universals) for their identity. These universals, in turn, depend for their existence upon their actual instantiations or the concrete potentialities for their instantiation.

One of the canonical advantages of Aristotelian universals over Platonic ones was that it did not need to commit to non-spatiotemporally located entities and thus did not violate ontological naturalism. Vetter's theory cannot, we have argued, enjoy such a benefit. While this does not amount to a knock-down argument against Vetter's position, we struggle to see what other good reasons there are to retain the commitment to **PPI**. Here's one hypothesis: Vetter (2015: 270) does not wish to attribute potentialities to Platonic universals (e.g. the potentiality to be instantiated), because "the picture of metaphysical modality that I am offering derives its attraction […] from the fact that it locates modality in properties, viz. potentialities, but also from the fact that it thereby locates modality in *objects* of the ordinary kind: concrete objects" (2015: 270). Perhaps Vetter is worried that, were one to adopt Platonic universals, she would be

drawn to attributing potentialities to them, rather than concrete objects. Note that the problem with this move is not that of attributing potentialities to abstract entities (since she attributes them to numbers: see Vetter 2015: 279-80) but rather that it tempts us into a 'catch-all solution', in which all we need to ground modal truths are the potentialities to be instantiated of all the universals. We think that the Platonist can resist such temptation just as well as the Aristotelian does. The point is orthogonal to the dependency of universals: it concerns the *bearers* of potentialities, not what potential properties are or what they depend upon. The Platonist is free to maintain that potentialities can only be instantiated by concrete objects, and yet the properties they are directed toward are metaphysically independent. Conversely, a proponent of Aristotelian universals is free to think that universals can be the bearers of potentialities: potentialities are had by non-fundamental entities, too.

More interestingly, resisting the 'catch-all' strategy does not mean that we can *never* attribute potentialities to universals – we think that it is open to the Platonist who accepts super-aliens to do so, in order to ground certain scientifically interesting truths. More precisely, we will try to show in the following section that Platonism is able to elegantly accommodate scientific possibilities involving so-called super-alien scenarios. In contrast, these are possibilities that Vetter's metaphysical framework struggles to accommodate. To reiterate: Platonists need not be committed to super-aliens, but we believe that such a commitment is highly beneficial to our understanding of scientific practice. In short, we think there are reasons to prefer a metaphysical outlook that is compatible with the possibility of super-aliens.

3. *Platonism, idealisation, and scientific possibility*

In the previous section we saw how Platonism is generally regarded as a non-naturalistic position, given that it allows entities which need not have concrete being. However, we do not think it follows that the theory of Platonism is divorced from science. Indeed, we think that certain kinds of scientific theorising may lend support to Platonism. For example, for various reasons scientists often reason about scenarios that involve what Vetter would regard as 'super-alien' properties, as defined above. If there are truths about which super-alien scenarios are and are not possible, then it is natural to enquire about the truthmakers for such claims. Here, the Platonists have a straightforward truthmaking story to tell: truths about super-alien possibilities are grounded in the modal profiles of uninstantiated properties. As we shall see, Vetter does not have the metaphysical resources to provide truthmakers for super-alien possibilities, which by her own admission leads her to deny that there are any

super-alien possibilities (2015: 270). Hence, if we accept Vetter's naturalistic framework, we can no longer take scientific talk of such possibilities at face value, which, as we shall see, leads to a disunified picture of scientific discourse about alien possibilities. We believe that this is a serious cost of Vetter's theory. To be clear, we do not think that, by itself, this problem is fatal. However, once the arguments of the previous section are also taken into consideration, the case for Platonism begins to look strong.

Why then does Vetter's potentiality-based theory of modality present a disunified picture of alien possibility? On the one hand, in cases where scientists theorise about possibilities concerning uninstantiated properties that some concrete thing has the potentiality to have (or produce), Vetter is happy to say that there are truths about such possibilities and that the relevant potentialities are the truthmakers. Indeed, science is awash with such truths, as exemplified by the *Journal of Computer-Aided Molecular Design*.[14] But on the other hand, in the case of alleged truths about super-alien possibilities – alleged possibilities that no concrete thing has the potentiality to realize – Vetter has to deny that there are any such truths, metaphysically speaking. Since, by definition, such possibilities are not tied to potentialities that some concrete thing has, Vetter has to deny that there are such possibilities after all. Super-alien possibilities might be assertible in the sense that they are epistemically possible, or true in some fictional sense, but strictly speaking no claims about the metaphysical possibility of super-alien properties can be (non-vacuously) true, since there is nothing in Vetter's framework that can serve as truthmaker. To be fair to Vetter, she accepts that this is a bullet she has to bite (2015: 269). However, we think that this is a more serious problem than Vetter acknowledges. Talk of super-alien possibilities is widespread in science. Moreover, as we shall see, some of the accepted laws of nature plausibly concern how *idealised* systems would behave in various circumstances and in many cases there are reasons for thinking that such systems are not physically realizable by anything. Hence, it seems that these laws – which are posited in science for various explanatory purposes – concern metaphysical possibilities that are super-alien on Vetter's definition. Given that we think laws of nature should be taken metaphysically seriously, we believe that super-alien possibilities should be too.

What then are examples of scientific claims about super-alien possibilities, and why do scientists think that these possibilities are important? Super-alien possibilities are typically expressed using what philosophers call counternomic or counterlegal conditionals. These are counterfactuals whose antecedents describe a scenario that is not possible in physical worlds like ours. The counter-

[14] We learnt about this journal from James Franklin (2015).

nomics that are of most interest in science involve idealisations. In such cases, the antecedent describes a scenario that is similar in some respects to a phenomenon that we are interested in in the real world, but which differs in other ways. In other words, such a description is an inaccurate representation of the real world phenomenon that we are interested in. It is inaccurate in the sense that it ignores certain properties that the real world system has or attributes properties that the real system does not have (see Psillos 2011: 7). Why then are counternomics useful in science, if they do not accurately represent scenarios in the real world? The answer is that idealisations help us to reason about the causal contributions made by specific properties of a system – contributions that are described in the consequent of the counternomic claim. Idealisations are "simplifying distortions" (Teller, 2012: 272), which make "immensely difficult physical problems computationally tractable and calculable to close approximations of their actual values" (Tan 2019: 44). If we want to investigate the relationship between, say, the length and trajectory of a natural pendulum, it seems we have no choice but to consider how a pendulum would behave in the absence of other causally relevant factors such as the mass of the pendulum string and air resistance. But since massless pendulum strings are not nomically possible, such reasoning inevitably rests upon counternomic claims.

Other examples of counternomic idealisations in science abound. Classical mechanics provides a rich stock of examples in the literature on counternomics, such as frictionless planes, point mass planets, ideal gases, models of projectile motion and so on. It is also plausible that modern theories in fundamental physics are heavily idealised. For example, "the quanta of quantum field theory are an artifact of describing space-time as flat" (Teller 2012: 269) and in physical theory "the velocity of light is the constant c (which of course we do not know precisely) – but only in a vacuum, and there are no perfect vacua" (Teller 2012: 263).

In any case, what is important for our purposes is that i) counternomic reasoning is employed in at least some areas of natural science, ii) many such counternomics describe properties that are super-alien in the sense defined earlier, and iii) some counternomic claims are true in a non-vacuous, metaphysically serious sense. We do not think that ii) is difficult to establish. It seems clear that nothing in our physical world has or will have the potentiality to instantiate or produce a frictionless plane, a massless piece of string, perfect vacua and so on. For example, in order for something to have the potentiality to produce a frictionless plane, it would have to be able to violate laws concerning molecular force. In favour of iii), it is prima facie plausible that there is an important difference between vacuously true counterpossible claims, whose antecedent describes a logically impossible scenario, and some of the counternomic claims

described above. In the case of vacuous counterpossibles, like "if 2+2 were to equal 5, then the moon would be made of cheese", any consequent whatsoever follows from the antecedent. But in the cases described above, there is surely a substantive right or wrong answer as to what would follow counterfactually if, say, my car were travelling on a frictionless plane. The non-vacuity of such counterfactuals becomes even more clear when we note that many idealised counterfactuals are used to capture facts about the actual laws of nature are in our world. Brian Ellis (1987: 54) has discussed a number of such examples. For instance, the principles of special relativity tell us how things would behave in inertial systems, though general relativity implies that nothing actually has the potential to physically realize such systems. The laws of thermodynamics provide other examples. Some of them describe how perfectly reversible heat engines would behave, even though other thermodynamic principles rule out the physical possibility of anything constituting or producing such engines. In short, if there are no super-alien possibilities, metaphysically speaking, then the metaphysical status of even scientifically supported laws of nature is brought into question. Assuming, then, that many counternomics in science are true in a non-vacuous, substantive way, it is natural to enquire about their truthmakers. The Platonist potentiality theorist has a ready-made answer: these counternomics are made true by the super-alien properties involved, such as *being a frictionless plane* or *being a Carnot engine*, which have a dispositional essence but which remain uninstantiated in our physical world.

What, then, is the problem for Vetter? As we saw earlier, Vetter does not have the metaphysical resources to provide truthmakers for counternomic claims, because all metaphysical possibilities have to be grounded in the potentialities of actual concrete things. Since no concrete thing has the potentiality to instantiate or produce an instantiation of a super-alien property, then there can be no such possibilities, metaphysically speaking. This means that scientific talk of super-alien possibilities must in some sense be second rate when compared with non-super alien possibilities. On Vetter's theory, non-super alien possibilities are in perfectly good standing: claims about them are non-vacuous and metaphysically substantive. But on Vetter's theory, as soon as we start thinking about super-alien possibilities, we are no longer thinking about genuine possibilities and are deceived if we think we are. This leaves us with a disunified account of scientific modal discourse, and we think this is a significant cost given the ubiquity of super-alien counternomic idealisation in science.

How then could someone like Vetter get around this worry? One option would be to accept that counternomics are true, but only in a vacuous sense. For reasons given above, we do not think this is a feasible option. Those who think otherwise are guilty of not paying close enough attention to examples

from science (see Tan 2019 for a detailed argument along these lines). A more promising solution would be to accept that our world contains objects with what Jenkins and Nolan (2012) call "impossible dispositions", which are dispositions for impossible manifestations (or stimuli). This solution allows someone like Vetter to maintain that super-alien manifestations are metaphysically impossible and at the same time the impossible dispositions would provide substantive truthmakers for the relevant counternomic claims.

Despite the promise of the impossible dispositions strategy, it is not one that Vetter herself endorses. In a discussion of the paper by Jenkins and Nolan, Vetter (2015: 250-257) discusses precisely this solution and rejects it. This is not surprising, given that the whole point of Vetter's project is to align possibility with potentiality; impossibilities arise when there is a lack of potentiality rather than a potentiality for an impossibility.[15] Many of Vetter's arguments against Jenkins and Nolan rely on a rejection of what she calls the "conditional conception" of dispositions (2015: Ch. 3). We do not have the space to discuss the details of Vetter's arguments but we agree with the conclusion that the impossible dispositions approach will not be plausible in many scientific cases. To be fair to Jenkins and Nolan, we do not think it is obviously absurd to say that an actual car has the disposition to move in a certain way on an inclined frictionless plane. However, we agree with Vetter (2015: 256) that in the example of the scientific idealisation that Jenkins and Nolan discuss (Jenkins and Nolan: 2012: 746), it is implausible to think there as an impossible disposition. The case allegedly requires us to say that a rabbit population has the disposition to increase by 0.1 rabbit per month. But given that these kinds of increases are ascribed to actual populations by scientists, surely such ascriptions must either be understood in non-literal way or else regarded as false, strictly speaking; in which case the dispositions for such increases can also be treated in a similar way (Vetter 2015: 256). Consider also the kind of counternomic claim discussed by Handfield: "if gravity had obeyed an inverse cube law, the planets would have had very different orbits" (2004: 403). We think it is implausible to suppose that masses have an impossible disposition to give rise to inverse cube gravitational behaviour. In our world, gravity is generated by mass, whose dispositional essence concerns inverse square behaviour rather than inverse cube behaviour. Moreover, if we are to accept that masses have impossible 'inverse cube' dispositions, we would surely have to posit an endless number of such

[15] This is also why Vetter should not appeal to the notion of masked potentialities in order to accommodate super-alien possibilities. For example, if we claimed that a heat engine had the potentiality for fully reversible behaviour, but that this potentiality would always be blocked by other properties of the engine, this would be tantamount to saying that the engine has a nomically impossible potentiality.

dispositions, with each one corresponding to a different inverse function. For these reasons, we believe that a Platonic analysis of Handfield's counternomic conditional is much more plausible. In line with Handfield's suggestion (2004: 406), we can interpret the antecedent as describing a case in which 'schmass' rather than mass in instantiated, where schmass's dispositional essence is to give rise to inverse cube behaviour. According to the Platonic analysis, the truthmaker for this counternomic is the property of schmass, which exists in an uninstantiated state. To be clear, this does not mean that it is possible that gravity obeyed an inverse cube law – only that super-aliens allow us to explain why that counternomic is true, as opposed to "if gravity had obeyed an inverse cube law, the planets would have had the very same orbits".

Another non-Platonic option to for someone like Vetter to take seriously is to say that although no particular object has the potentiality to produce super-alien instantiations, the world as a whole does. Indeed, Vetter discusses this strategy and explores whether it could be used as a catch-all solution for alleged cases of possibility for which potentialities are not easy to find (2015: 257-263). [16] As Vetter points out (2015: 261), this general strategy will remind some of the explanation of conservation laws given by Bigelow, Ellis and Lierse (1992), according to which conservation principles are grounded in the dispositional essence of the 'world kind'. Vetter is however sceptical of this solution and argues convincingly that spelling out the nature of the world-level object in the appropriate way is difficult to do (2015: 258-261). We would merely add that ascribing potentialities to the world as a whole also leaves us with explanations which arguably suffer from the same problems as those provided by Bigelow, Ellis and Lierse in the case of conservation laws. As Livanios (2010: 302) argues, such explanations are either poor ones or deeply *ad hoc*. According to Livanios such explanations are often poor because they are too coarse grained. That is to say, world-level explanations are so general in character that they are not informative. Moreover, as Tugby discusses elsewhere (2017: 2071), the world – level strategy is suspiciously easy to employ, because it provides an automatic recipe for explaining any modal phenomenon that is otherwise difficult to explain. Surely doing modal metaphysics should be more difficult than this.

What then are the remaining options for those who accept Vetter's metaphysical framework? We think that the least problematic options all involve biting the bullet, by conceding that there are no super-alien metaphysical possibilities. There are at least two ways of doing this. One is to employ a fictionalist account of scientific idealisation along the lines Frigg's theory of scientific

[16] We are also grateful to Daniel Nolan for raising this possibility during discussion.

models (2010). According to such a theory, claims about counternomics are not literally true but are true in some weaker sense: true according to scientific fiction. The other option is one that Vetter has recently endorsed (2016), which involves viewing claims about super-alien possibilities as counterpossibles that can be true, but which concern *epistemic* rather than metaphysical possibility. On this account, true counternomic claims are truths about the compatibility of a certain counterfactual (or counterpossible) with our knowledge or evidence. Claims about the potentialities or dispositions of things are, in contrast, metaphysical or "circumstantial" (2016: 2694). Unfortunately, we do not have the space to discuss the proposals of Frigg and Vetter in the detail that they deserve. We accept that if Platonism is rejected, then the fictionalist and epistemic approaches will be strong contenders. However, our point for current purposes is that, unlike Platonism, both theories incur the cost of leaving us with a disunified treatment of scientific modal discourse. On these theories, many scientific modal claims are both literally true and metaphysically serious, but as soon as scientists engage in idealised, counternomic modal discourse, the claims are either literally false or not metaphysically serious. We cannot help feeling that this makes the latter sort of scientific discourse appear second rate. Moreover, if people like Teller (2012) are right, and almost all fundamental scientific theories are idealised and rest heavily on counternomic assumptions, then we worry that the fictionalist and epistemic approaches strike a blow against scientific realism itself. If one wants to take scientific theories metaphysically seriously, then Platonism is a very tempting route to take.

4. *Conclusions*

We have argued that Vetter cannot accommodate both Hardcore Actualism and Naturalism. We recommend that Vetter loosens her commitment to Naturalism and accepts a Platonic theory of potentiality or something like it. The Platonic theory provides an elegant solution to the fundamental puzzle of potentiality and brings a variety of peripheral modal benefits, especially in the context of natural science.

Giacomo Giannini
giacomo.giannini@durham.ac.uk
PhD student, Durham University

Matthew Tugby
matthew.tugby@durham.ac.uk
Associate Professor, Durham University

Acknowledgments

Matthew Tugby would like to thank the organisers and participants at the conference 'Counterpossibles, Counternomics and Causal Theories of Properties', held at the University of Cologne in April 2019, where we received valuable feedback on the material about counternomics. The conference was funded by the DFG (German Research Foundation) research group FOR 2495 Inductive Metaphysics.

Giacomo Giannini carried out the research on which this paper is based while being supported by the AHRC's Northern Bridge Doctoral Training Partnership.

References

Armstrong, D.M., 1978, *Universals and Scientific Realism*, Vols. I and I,. Cambridge University Press, Cambridge.

Armstrong, D.M., 1981, *The Nature of Mind and Other Essays*, Cornell University Press, Ithaca.

Armstrong, D.M., 1989, *Universals: An Opinionated Introduction*, Westview, Boulder, CO.

Armstrong, D.M., 1997, *A World of States of Affairs.* Cambridge University Press, Cambridge.

Azzouni, J., 2004, *Deflating Existential Consequence: A Case for Nominalism*, Oxford University Press, New York.

Barker, S., 2009, "Dispositional monism, relational constitution and quiddities", in *Analysis*, 69: 242-250.

Bigelow, J., Ellis, B., and Lierse, C., 1992, "The world as one of a kind: natural necessity and laws of nature", in *The British Journal for the Philosophy of Science*, 43: 371-388.

Bird, A., 1998, "Dispositions and antidotes", in *The Philosophical Quarterly*, 48: 227-234.

Bird, A., 2006, "Potency and modality", in *Synthese*, 149: 491-508.

Bird, A., 2007, *Nature's Metaphysics: Laws and Properties*, Oxford University Press, Oxford.

Bird, A., 2016, "Overpowering: How the Powers Ontology Has Overreached Itself", in *Mind*, 125: 341-383.

Borghini, A., Williams, N.E., 2008, "A dispositional theory of possibility", *dialectica*, 62: 21-41.

Contessa, G., 2008, "Modal Truthmakers and Two Varieties of Actualism", in *Synthese*, 174: 341-353.

Contessa, G., 2019, "Powerful Qualities or Pure Powers?", *Metaphysica*, 20 (1): 5-33.

deRosset, L., 2010, "Getting priority straight", in *Philosophical Studies*, 149: 73-97.

Ellis, B., 1987, "The ontology of scientific realism", in P. Pettit, R. Sylvan and J. Norman (eds.), *Metaphysics and Morality: Essays in Honour of J.J.C. Smart*, Blackwell, Blackwell: 50-70.

Fine, K., 1994, "Essence and modality", in J. Tomberlin (ed.), *Philosophical Perspectives 8: Logic and Language*, Ridgeview, Atascadero CA: 1-16.

Fine, K., 2015, "Unified foundations for essence and ground", in *Journal of the American Philosophical Association*, 1: 296-311.

Franklin, J., 2015, "Uninstantiated properties and semi-Platonist Aristotelianism", in *Review of Metaphysics*, 69, 25-45.

Frigg, R., 2010, "Models and fiction", in *Synthese*, 172: 251-268.

Giannotti, J., forthcoming, "The identity theory of powers revisited", in *Erkenntnis*.

Handfield, T., 2004, "Counterlegals and necessary laws", in *The Philosophical Quarterly*, 54: 402-419.

Heil, J., 2003, *From an Ontological Point of View*, Oxford University Press, Oxford.

Heil, J., 2012, *The Universe as We Find It*, Oxford University Press, Oxford.

Jaag, S., 2014, "Dispositional essentialism and the grounding of natural modality", in *Philosophers' Imprint*, 14.

Jacobs, J., 2011, "Powerful qualities, not pure powers", in *The Monist*, 94 (1): 81-102.

Jenkins, C. and Nolan, Daniel, 2012, "Disposition impossible", in *Noûs*, 46: 732-53.

Ingthorsson, R., 2013, "Properties: Qualities, Powers, or Both?", *Dialectica*, 67: 55-80.

Ingthorsson, R., 2015, "The Regress of Pure Powers Revisited", in *European Journal of Philosophy*, 23: 529-541

Lewis, D., 1986, *On the Plurality of Worlds*, Blackwell, Oxford.

Livanios, V., 2010, "Symmetries, dispositions and essences", in *Philosophical Studies*, 148: 295-305.

Lowe, E.J., 2010, "On the individuation of powers", in A. Marmodoro (ed.), *The Metaphysics of Powers: Their Grounding and Their Manifestations*, Routledge, Abingdon: 8-26.

Kripke, S., 1980, *Naming and Necessity*, Harvard University Press, Cambridge MA.

Manley, D., and Wasserman, Ryan, 2008, "On linking dispositions and conditionals", in *Mind*, 117: 59-84.

Martin, C. B., 1994, "Dispositions and conditionals", in *Philosophical Quarterly*, 44: 1-8.

Martin, C. B., 1997, "On the Need for Properties: The Road to Pythagoreanism and Back", in *Synthese*, 112: 193-231.

Martin, C. B., 2008, *The Mind in Nature*, Oxford University Press, Oxford.

Molnar, G., 2003, *Powers: A Study in Metaphysics*, edited by S. Mumford, Oxford University Press, Oxford.

Mumford, S., 2004, *Laws in Nature*, Routledge, London.

Mumford, S., and Anjum, Rani L., 2011, *Getting Causes from Powers*, Oxford University Press, Oxford.

Plantinga, A., 1983, *The Nature of Necessity*, Oxford University Press, Oxford.

Priest, G., 2005, *Towards Non-Being. The Logic and Metaphysics of Intentionality*, Oxford University Press, Oxford.

Psillos, S., 2011, "Living with the abstract: realism and models", in *Synthese*, 180: 3-17

Rosen, G., 2010, "Metaphysical dependence: grounding and reduction", in B. Hale and A. Hoffman (eds.), *Modality: Metaphysics, Logic, and Epistemology*, Oxford University Press, Oxford.

Sider, T., 2011, *Writing the Book of the World.* Oxford University Press, Oxford.

Strawson, G., 2008, "The identity of the categorical and the dispositional", in *Analysis*, 68, 271-282.

Tan, P., 2019, "Counterpossible non-vacuity in scientific practice", in *The Journal of Philosophy*, 116: 32-60.

Taylor, H., 2013, "In defence of powerful qualities", in *Metaphysica*, 14: 93-107.

Taylor, H., 2018, "Powerful qualities and pure powers", in *Philosophical Studies*, 175: 1423-1440,

Teller, P., 2012, "Modeling, truth and philosophy", in *Metaphilosophy*, 43: 257-274.

Tugby, M., 2012, "The metaphysics of pan-dispositionalism", in A. Bird, B. Ellis, and H. Sankey (eds.), *Properties, Powers and Structures: Issues in the Metaphysics of Realism*, Routledge, Abingdon: 165-179.

Tugby, M., 2013, "Platonic dispositionalism", in *Mind*, 122: 451-480.

Tugby, M., 2016, "Universals, laws and governance", in *Philosophical Studies*, 173: 1147-1163.

Tugby, M., 2017, "The problem of retention", in *Synthese*, 194: 2053-2075.

Vetter, B., 2015, *Potentiality: From Dispositions to Modality*, Oxford University Press, Oxford.

Vetter, B., 2016, "Counterpossibles (not only) for dispositionalists", in *Philosophical Studies*, 173: 2681-2700.

Vetter, B., 2018, "A plenitude of powers", in *Synthese*, DOI 10.1007/s11229-018-1758-z

Williamson, T., 2002, "Necessary existents" in A. O'Hear (ed.), *Royal Institute of Philosophy Supplement*, Cambridge University Press, Cambridge: 269-287.

Williamson, T., 2013, *Modal Logic as Metaphysics*, Oxford University Press, Oxford.

Potentialities as properties

Jennifer McKitrick

Abstract: In *Potentiality: From Dispositions to Modality*, Barbara Vetter attempts to ground modality in properties that she calls "potentialities". Whether potentialities are up to this task depends on what *properties* are. However, major accounts of the metaphysics of properties, such as Class Nominalism, Trope Theory, Immanent Realism, Platonism, are incompatible with Vetter's claims about potentialities. Nevertheless, a modified account of potentialities might be compatible with Immanent Realism or Trope Nominalism.

Keywords: disposition; potentiality; power; property; universal; trope; nominalism.

1. *Introduction*

In *Potentiality: From Dispositions to Modality*, Barbara Vetter attempts to ground modality in dispositional properties that she calls "potentialities".[1] Potentialities are said to have a number of features intended to make them suitable for this job. Seeing as potentialities are properties, whether they can do this work depends, in part, on what properties are. While Vetter makes a few suggestive remarks in her book, she gives little consideration to the question of which theory of properties is compatible with her claims about potentialities. In this paper, I explore a worry that no major approach to understanding properties is compatible with the claims that Vetter makes in *Potentiality*.[2] I suggest a modified account of potentialities that might be compatible with Immanent Realism or Trope Nominalism, then close by considering two objections to my proposal.

2. *Vetter's Account of Dispositions and Potentialities*

Vetter introduces potentialities by way of the more familiar notion of dispositions – properties like fragility and solubility. On the standard conception

[1] Others use "potentiality" in other ways. Here, I focus on Vetter's stipulated use of the term, and make no claims about potentiality in general.

[2] This chapter expands on points that I made in McKitrick 2019.

philinq VIII, 1-2020, pp. 141-166
ISSN (print) 2281-8618-ETS DOI: 10.4454/philinq.v8i1.279

of dispositions, they are associated with conditionals. This association can be as strong as defining dispositions in terms of conditionals. That is to say, for example, "x is fragile" means (roughly) "if x were struck, x would break". According to a simple conditional analysis:

"x has a disposition to manifest M in a certain stimulus condition C" $=_{df}$
"If x were in C, then x would manifest M".

As many have before her,[3] Vetter rejects the semantic reduction of disposition ascriptions to conditional statements. Manley and Wasserman are among the other theorists who also argue against standard conditional analyses (Manley and Wasserman 2007). They point out that some dispositions such as loquaciousness seem to have no stimulus conditions. In those cases, it is not clear what the content of the antecedent of the associated conditional could be.

In addition to the missing stimulus issue, Manley and Wasserman make three interrelated observations about disposition ascriptions which Vetter picks up on. First, disposition ascriptions are context-sensitive.[4] A wooden beam that is called "fragile" on a construction site would not be called "fragile" in an antique shop. Any analysis that specifies one stimulus-manifestation pair to correspond to "fragile" in one context renders the wrong result in other contexts. Second, disposition ascriptions can be comparative. A wine glass is more fragile than a coffee mug. Any analysis that associates the same conditional with the wine glass's fragility and the mug's fragility cannot do justice to such assertions. Third, the applicability of a dispositional predicate can be a matter of degree because they are "gradable". Some things are "extremely fragile" while other things are "somewhat fragile". Conditional statements, on the other hand, do not admit of degrees of truth (on most standard logics). If an analyst were to translate all disposition ascriptions in terms of pairs of circumstances and manifestations, they would have to find a pair for every degree of fragility, as well as pairs for every degree of every other gradable dispositional predicate.

In light of such problems, Manley, Wasserman and others argue that dispositions are connected to conditionals in more complex ways.[5] Vetter, on the other hand, rejects the association between dispositions and conditionals entirely. Instead, she claims that disposition ascriptions are possibility statements. On her account "x is breakable" means roughly "x can break", with no stimulus specified. While Manley and Wasserman propose a Proportionality

[3] For example, see Martin 1994 and Bird 1998.

[4] There are dispositional theories that hold that dispositions are context-independent properties (e.g. Mumford and Anjum 2011).

[5] For example, see Lewis 1997; Bird 2007; Contessa 2013; McKitrick 2018.

Account according to which disposition ascriptions implicitly reference a contextually determined adequate proportion of possible circumstances, Vetter looks for mind-independent, context-neutral properties as the truth-makers for such statements. According to Vetter, if someone thought that context-sensitive terms such as "fragile" denote properties, then

> they would have to countenance context-sensitive properties. But reality is not context-sensitive. Context-sensitivity is a matter of language, not the world. Context-sensitive expressions receive different semantic values in different contexts of utterance; reality provides the semantic values […] (Vetter 2015: 80)

For Vetter, the metaphysical background in virtue of which disposition ascriptions are true is constituted by potentialities. Potentialities are disposition-like properties that are individuated by their manifestations. They are irreducible and metaphysically basic: While some potentialities are grounded in other potentialities "as we progress from the less to the more fundamental levels, we will always find potentialities. It's potentiality 'all the way down'" (25).[6] For example, salt is soluble because its sodium and chlorine atoms are held together by ionic bonds, and being ionically bonded is grounded in electric charge, so arguably solubility is grounded in electric charge. If electric charge is absolutely fundamental, then it is an irreducible and metaphysically basic potentiality. But even if it is not fundamental, it is grounded in further potentialities, on Vetter's view. While Vetter refrains from making claims about absolute fundamentality, the view the she is clearly rejecting is one in which something has a potentiality in virtue of other facts that do not involve potentialities. There is no kind of ontological entity more fundamental than potentiality, for Vetter.

In addition to being context-independent, potentialities differ from dispositions in a number of other ways. Ordinary disposition ascriptions typically suggest that the occurrence of the manifestation is highly likely, but not certain. For example, while it is not metaphysically impossible that I become a drug dealer, I would not say that I am "disposed" to become a drug dealer.[7] Furthermore, while I may be necessarily self-identical, I would not say that I am "disposed" to be self-identical. Potentialities, on the other hand, come in a wide spectrum of degrees, meaning that the likelihood of their manifestations ranges from not impossible to necessary (21). The part of the spectrum where dispositional ascriptions are usually appropriate typically ranges over the upper half, but short of necessity. This picture nicely accommodates the fact that dispositional predicates are gradable. Because a potentiality comes in

[6] Numbers in brackets, unless otherwise specified, refer to *Potentiality*.

[7] Thanks to Edward Becker for this example.

a wide spectrum of degrees, it is a determinable property and all of the specific degrees to which it can be possessed are its determinates (43, 95). A sturdy crowbar has some degree of breakability because it is not unbreakable. A fragile glass and a sturdy crowbar have different determinates of the determinable potentiality breakability.

Intuitively, if specific dispositions are determinates of determinable potentialities, and potentialities are more fundamental than specific dispositions, then potentialities are determinables which are more fundamental than their determinates. Several statements from Potentiality demonstrate Vetter's commitment to this view:

1. "*Potentiality comes in degrees*" (emphasis mine, 95).
2. *A potentiality is a determinable and its degrees are its determinates*: "We can think of that relation [between a given potentiality and its various degrees] again, on the model of the relation between height and the particular heights: that is, as a relation between a determinable and its determinates" (95).
3. *A disposition is a degree of a potentiality*: "having a disposition such as fragility is a matter of having the right potentiality (in this case the potentiality to break or be broken) to a contextually sufficient degree" (22).
4. *Therefore, a disposition is a determinate of a determinable potentiality.*
5. *Potentialities ground dispositions*: "The notion of a potentiality has been introduced as the metaphysical background to the context-dependent notion of a disposition" (96).
6. *The dispositions which are grounded by a potentiality include the dispositions which are degrees of that potentiality*: "having a disposition such as fragility *is a matter of* having the right potentiality (in this case the potentiality to break or be broken) to a contextually sufficient degree" (my emphasis, 22).
7. *If A grounds B, then A is more fundamental than B*: "the more fundamental grounds the less fundamental" (23).
8. *Therefore, determinable potentialities are more fundamental than their determinate dispositions*: "the general dispositions are not only equally fundamental as the specific ones, they are *more* fundamental" (57).[8]

While I will take issue with this conclusion later, I do not take this argument to be a *reductio ad absurdum* of Vetter's view, but merely an explication of what is implicit in the text which I believe she would endorse. Other remarks are

[8] It is not clear from the context of this quote whether it applies to determinables and their determinates as well as single and multi-track dispositions, and so the rest of the argument above is necessary to establish this conclusion.

also suggestive. She asks rhetorically "why should the maximal determinate potentiality be more natural than its determinable? […] Simply stipulating the one to be more natural than the other is ad hoc" (287). (Note that "more natural than" tracks "more fundamental than" as well as "grounds" (29).)

3. *Theories of Properties*

3.1. Vetter's Remarks

Since Vetter's book *Potentiality* does not include a lengthy discussion of the metaphysics of properties, we can quickly examine most of what is written there on that topic. In articulating her background assumptions, Vetter writes:

> I will be maximally liberal about which properties there are. The properties that there are include the properties of: being electrically charged and being fragile, being green and being grue, being self-identical and being identical to me, and being such that grass is green. Some of these properties, to be sure, are more natural than others. […] In general, possession of the more natural properties grounds possession of the less natural ones, thus we might equally speak of more or less fundamental properties. […] I do not take the fact that a property's instantiation is grounded in the instantiation of other properties as a reason to reject the grounded property – on the contrary. (29)

She goes on to speculate that her claims about non-natural properties can be reformulated by those with more austere preferences (30). Consequently, this non-committal liberalism about properties places few constraints on which metaphysics of properties that she can employ. However, Vetter makes other remarks that are more suggestive:

> I assume that there are properties, and that a property is the kind of entity that different particulars can share. When two apples are both red, then there is a property that they both possess, the property of being red. Accordingly, different things can share a potentiality for the same property: the sheets of paper in this book all share the potential to burn. (29)

Furthermore, she assumes that "there are properties, in the sense of universals that can be multiply instantiated" (270). I'll call this view "Aristotelian" or "Immanent" Realism, according to which two things that have the same property literally share a single entity – the universal.[9] It contrasts with a Platonic or

[9] While I follow Vetter in calling this view "Aristotelian", I leave it to scholars of Aristotle to debate whether Immanent Realism is accurately attributable to Aristotle.

Transcendent Realism according to objects instantiate properties by standing in relations to universals that exist elsewhere. Vetter resists Platonism, writing "[a] Platonist approach […] goes against the motivation of this book: the idea that modality is ultimately a matter of how objects, in the ordinary, concrete sense, are" (270). However, like her liberalism about properties, Vetter regards her preference for Immanent Realism as a non-issue for the purposes of her current project. She speculates that her theory of potentialities is consistent with certain forms of Nominalism, according to which universals do not exist. She writes:

> Nominalists about properties have different strategies in dealing with our ubiquitous talk of properties shared by things. One is to claim that when we say that two apples share a property, what we say is not in fact true, but something similar is (the two red apples have more or less exactly resembling tropes, or they are both members of the same class of particulars that resemble each other in a certain respect). Another is to claim that what we say is true, but analyse away the apparent appeal to shared properties (what we really say is that things have resembling tropes, or that they belong to the same class). I suspect that both strategies could be applied throughout the book, making only slight alterations to my central claims and arguments […] But at the moment, I have no more than that suspicion to offer to the nominalist. (29)

My goal is to scrutinize Vetter's suspicion, and whether any familiar account of properties is consistent with her theory of potentialities.

3.2. Nominalism

According to Nominalists, either properties can be accounted for without appealing to universals, or properties need not be accounted for at all. According to Vetter, Nominalists say that claims about objects sharing properties are ways to talk about something else – the talk of shared properties is, as she says, "analyzed away". So, if potentialities are properties, Nominalists must analyze away talk of shared potentialities. It is not clear how this could be consistent with Vetter's claim: "I assume non-reductive realism about potentiality […]. dispositions, or potentialities, are metaphysically basic, primitive, irreducible" (24), and giving up on this assumption strikes me as more than just a slight alteration of her central claims. Let's look at a few varieties of Nominalism in more detail. [10]

According to Predicate Nominalism, two objects share the same property if the same predicate applies to them. According to Concept Nominalism, two

[10] For a survey of varieties of nominalism, see Armstrong 1978 and Rodriguez-Pereyra 2016.

objects share the same property if they fall under the same concept. On either view, there is nothing about the objects in virtue of which this is so, since that would suggest that the objects share a property, and such claims are to be reduced to claims about predicates or concepts. One problem for Vetter's realist approach to metaphysics is that concepts and predicates are part of contingent human psychological and linguistic practices which vary across time and place. Talk of "merely possible" predicates or concepts might seem to do some work, but remember that Vetter wants to ground possibility in potentialities of actual objects. On Vetter's view, for a predicate to be possible is for something to have a potentiality for that predicate to exist. If potentialities are properties, and properties are analyzed in terms of possible predicates, and possible predicates are analyzed in terms of potentialities, then the analysis might be problematically circular. Realism about potentialities is difficult to square with Predicate or Concept Nominalism.

According to Class Nominalism, properties are classes of objects. For an object to have a property is for it to be a member of a certain class. Again, there is nothing about the objects in virtue of which they are classed together, for that would suggest that they share a property, but that just means they are in the same class. According to David Lewis (1983), any class of possible objects, no matter how heterogeneous its members, is a property. Some properties are special ("natural") because the predicate "__is a natural class" applies to them. But since this predicate of classes cannot be analyzed in terms of shared properties, it is taken as primitive.

Vetter suggests that Class Nominalism is consistent with her views. However, there are problems with taking Vetter's potentialities to be natural classes of objects. If we take the Lewisian approach, classes are sets. Sets necessarily have the members that they do, and set membership is all or nothing, not a matter of degree. A wine glass is more fragile than a coffee mug, but it makes no sense to say that it has a greater degree of membership in the set of fragile things than the coffee mug does.

How could a Class Nominalist interpret Vetter's claim that "Like height, potentiality comes in degrees" (95)? Perhaps matters of degree can be explained in terms of subsets. Suppose that a determinable potentiality 'breakability' is a large set of all-but-unbreakable objects, and different determinate breakabilities are subsets of this large set. This doesn't get us gradability, since being in a set with other things that are equally breakable does not entail anything about a thing's breakability with respect to anything else. More work needs to be done by the Nominalist fan of potentialities to cash this out. One suggestion is that the super-set partitions naturally into certain subsets, and those subsets are ordered. This ordering of the subsets licenses a loose manner of talk of

ordered membership in their union (which is identical to the determinable property, on this suggestion).[11] However, while this sounds like a promising approach for a nominalist account of determinable quantities, it is in tension with Vetter's claim that potentialities are fundamental. One of the defining characteristics of potentialities, gradability, would turn out to be a merely loose manner of speaking about a relations between a potentiality's subsets. Alternatively, quantities can be grounded in proportionality relations, or explained in terms of certain relational predicates (Eddon 2013). But according to Maya Eddon, on such accounts "all facts about quantity are ultimately grounded in relations among objects" (Eddon 2013: 12). So, if potentialities were nominalist quantities, they would be grounded in relations among objects, and this would make those relations more fundamental than potentialities. This is inconsistent with Vetter's claim that potentialities are fundamental.

Furthermore, if Vetter adopted Class Nominalism, she would confront the challenge of distinguishing co-extensive properties. For example, if it turns out that everything that can possibly be broken is the same set as everything that can possibly burn, then breakability and inflammability would be the same potentiality. This familiar problem for Class Nominalism is especially salient for Vetter because a great many potentialities have extremely inclusive extensions: The only particulars that are outside of the extension of the 'potentiality to be F' are those particulars that are necessarily not-F. Furthermore, on Vetter's view, everything has the potential to be self-identical, to be such that triangles have three sides – to have any property that everything necessarily has. If Vetter embraced Class Nominalism as well, there would be but one necessarily possessed potentiality – the class of everything. So, 'the potentiality to be self-identical' would be identical to 'the potentiality to be such that triangles have three sides,' as well as all of the rest.

Some philosophers differentiate contingently co-extensive properties by appeal to merely possible objects (Lewis 1983). On this Lewisian approach, even if two properties happen to have the same extension in the actual world, they have different extensions across possible worlds, and this differentiates the properties. However, taking this approach would be problematic for Vetter. If a potentiality is a property, and a property is a set of possibilia, then potentialities are not well-suited to ground all of modality. Furthermore, Class Nominalism has trouble explaining causal powers. Having a potentiality means that an object can do things – that it is powerful. I doubt that a fan of potentialities would endorse the idea that being a member of class could enable an object to do anything that it couldn't already do.

[11] I owe this suggestion to Nick Jones.

Vetter also suggests that her view is compatible with Resemblance Nominalism – the view according to which properties are classes of objects that resemble one another. This view is similar to Class Nominalism in that it identifies properties with classes. Consequently, Vetter's difficulties for adopting Class Nominalism noted above apply here as well: It is unclear what it would mean for a class to come in degrees; there could be no distinct co-extensive potentialities, and class membership is a problematic source of causal power. To these problems, Resemblance Nominalism adds a few more. Famously, the Resemblance Nominalist cannot define resemblance in terms of shared properties, and thus has difficulty explaining what it means for objects to resemble each other "in a certain respect". Lacking an account of resemblance, Resemblance Nominalism cannot explain how determinate degrees of the same potentiality *imperfectly* resemble each other.

Another nominalist position that Vetter entertains is Trope Nominalism. Tropes are particular property instances, such as this page's whiteness. General properties are classes of similar tropes, such as the class of all particular whitenesses of the same shade. Similarity must be taken as primitive for reasons which should be familiar by now. Applying this theory to potentialities, there is no obvious problem with saying that a glass's specific degree of fragility is a determinate degree-of-potentiality trope. But recall, potentialities are determinable. If Vetter were to adopt Trope Nominalism, would she say that the same potentiality trope is both a determinate fragility trope and a determinable breakability trope? Or does the glass have determinable breakability in addition to its specific fragility? A single trope is not the kind of thing that comes in degrees, and for reasons similar to those raised above, it is not clear what it would mean for a set of tropes to come in degrees.

One possibility is that a set of similar fragility tropes is a property which is a determinate degree of breakability and a subset of determinable breakability. Then the glass's fragility trope would be both an instance of a determinate fragility and determinable breakability in virtue of being a member of both sets of tropes. But it is not clear what other tropes would be included in a determinable set, since Trope Theory does not explain imperfect resemblance. By analogy, we might want to include a scarlet trope and a crimson trope in the set of tropes that is the determinable general property 'redness,' but it is difficult to say what it is about the nature of the tropes that justifies this preference. In addition to "resemblance", perhaps the Trope Nominalist must take "imperfect resemblance" and "degrees of resemblance" as primitive as well. Finally, recall that determinable potentialities are supposed to be more fundamental than determinate degrees of potentiality. If determinable potentialities are sets of determinate potentialities tropes, then applying Trope Nominalism

to Vetter's theory of potentialities has the consequence that some sets are more fundamental than their members. Since Vetter assumes "that which grounds is more fundamental than that which is grounded in it" (27) she would be committed to the view that some sets ground their members. This runs counter to a common supposition that sets are grounded by their members (Correia and Schneider 2012: 1, 20).

One type of Nominalism that Vetter does not consider is Causal Nominalism, due to Ann Whittle (2009). Causal Nominalism might appear to be the variety of Nominalism most suited for potentialities, for it is the view that a property is a set of particulars that are similar with respect to causal role. As Whittle puts it, the property of F-ness is "the set of particulars all of which realize the functional role definitive of F-ness" (Whittle 2012: 248). F-ness, so defined, seems like a good candidate for being a potentiality. However, upon closer examination, Whittle's Causal Nominalism is not a good fit with Vetter's potentialities. As Whittle uses the term, "Nominalism" is the view that "everything that exists is particular" and "there are no basic property instances or tropes" (244). For Whittle, properties of any kind are reducible to particulars, and thus she provides "deflationary account of powers" (268). Consequently, this view is not compatible with Vetter's view that potentialities are fundamental and irreducible. Furthermore, Whittle's picture includes irreducible modal facts, specifically conditional or functional facts. In particular, when an object a has a disposition F "all there is to a's being F is that a particular cluster of causal conditionals holds true of a" (280). Whittle also writes:

> there are irreducible functional facts about what particulars can do. […] at the level of [perfectly] natural properties, a is F iff it could do X in circumstances C1 etc. – there is nothing further we can appeal to which accounts for the behaviour of the particulars in question. (Whittle: 2012: 283)

In other words, a set of particulars is a property if and only if certain irreducible modal facts hold of members of that set. The problem with applying Causal Nominalism to Vetter's view is that, if such sets of particulars were to constitute potentialities, then potentialities could not ground all of modality. This is because it would be viciously circular to use potentialities to ground the modal facts which determine which sets of objects constitute potentialities in the first place. While Whittle's particularly austere version of Nominalism is not central to Causal Nominalism (McKitrick 2018: 98-99), the essence of the view is that objects share the same property in virtue of their causal similarity. Insofar as causation is a modal notion, a potentiality cannot be a more fundamental modality if it is a property as defined by Causal Nominalism.

I do not claim to have conclusively shown that no version of Nominalism could be devised to be compatible with Vetter's claims about potentialities. However, I do hope to have shown that the most familiar forms of Nominalism are difficult to square with the claim that all of modality is grounded by properties which are irreducible potentialities.

3.3. Realism

Perhaps Vetter's suspicion that Nominalists could get on board with her project were too optimistic, and the theory of properties that she implicitly assumes throughout the work is more promising. Vetter writes:

> my preferred view of properties is what is often labelled the 'Aristotelian' view. I hold that there are properties, but that they are *in rebus*: their existence derives from how things are. The opposed, Platonist, picture has it that properties exist independently, and that objects are as they are in virtue of partaking in, or instantiating, the relevant universals. […] Perhaps the best known contemporary defender of an Aristotelian approach to universals is David Armstrong (271).

There are several things to note about this passage. First, in saying that properties are "*in rebus*", I take Vetter to be saying a property is "in the thing", and hence located where the thing – its instance – is. Since the same property can be shared by distinct particulars, it follows that properties are multiply-located. Second, she says that the view that she favors is defended by Armstrong, so I assume that Armstrong's work on properties is a relevant touchstone. Third, in saying that the existence of properties "derives from how things are" I take Vetter to be asserting the Immanent Realist's instantiation condition, according to which properties cannot exist without things that potentially instantiate them, not the Nominalist position that potentialities are mere derivative entities.

However, Vetter's preference for Immanent Realism runs up against her view that potentialities come in different degrees. It is not clear what it means for a universal to come in different degrees. A universal cannot have a greater or lesser degree than it does. So, it must be that a particular token or instance of the potentiality "has" more of the potentiality universal than another token of that potentiality. But if a universal is wholly present wherever it is instantiated, it cannot possibly differ across instances. If two things are similar because they literally share one thing, they must be perfectly similar in that respect. So if the fragile glass is highly breakable and the sturdy crowbar is less breakable, it is hard to see how that is explained in terms of their literally sharing one and the same 'breakability.'

Maybe the fragile glass and sturdy crowbar don't literally share the same 'breakability', but each have specific degrees of breakability, each of which are determinates of the determinable breakability. Analogously, a color comes in different shades, and particulars instantiate that color by instantiating one of the specific shades. This would cohere with Vetter's claim that potentialities are determinable, and specific degrees of potentiality are their determinates. If breakability is a potentiality universal, saying that it comes in degrees is a way of saying that there is an ordered set of universals, with the different "degrees" falling under the determinable potentiality. Vetter indicates this when she writes:

> I will sometimes speak of that potentiality as being possessed to a certain degree […] This is merely a more idiomatic way of expressing that an object has the determinable potentiality by having a given (degree-) determinate (95).

But saying that an object has a determinable property "by having a given (degree-) determinate" suggests that its having a determinable property *is grounded by* its having a certain determinate property. Given Vetter's acceptance of the idea that grounds are more fundamental than that which is grounded, if the determinate property grounds the determinable property, then the determinate property is more fundamental. This is inconsistent with Vetter's view that the determinable potentiality is more fundamental than its determinates.

Furthermore, this picture raises questions about property location. While it is reasonable for an Immanent Realist to say that the determinate properties are located in their instances, it is less clear where determinable properties are located. If the determinable 'breakability' were wholly present in each of its instances, it would have to be identical in each instance, but it is not clear that it is. Vetter seems implicitly aware of this issue when she writes:

> [A] criterion for naturalness that is often used is this: the more natural a property is, the more perfect the resemblance for which it makes. Perfectly natural, fundamental properties make for perfect resemblance. (Lewis (1983) appeals to this criterion in delineating the perfectly natural properties.) This criterion simply cuts no ice between [general] electric charge and [specific] charge e∗. Both, I take it, make for perfect resemblance. The former makes for more resemblance, but it is unclear whether that is relevant to its comparative degree of naturalness. (58fn14)

However, the criterion does seem to cut some ice between determinable breakability and a determinate degree of breakability. Breakability, which a crowbar and a wine glass share, makes for less similarity than a specific degree of fragility which certain types of glassware share. If properties that make for

more similarity are more natural and thus more fundamental, then fragility is more fundamental than breakability, contrary to Vetter's claims.

It might be helpful to look to the best known defender of Immanent Realism, and his treatment of determinables and imperfect resemblance. While Armstrong's views shift over time (1978; 1997; 2010), one of his long-standing positions is that particulars instantiate only *determinate* universals, and those instantiations are the truth-makers for attributions of determinable predicates (1978: 61). In other words, determinable predicates such as "breakable" do not designate universals. Obviously, this version of Immanent Realism is not an option for Vetter, for she holds that potentialities are real, determinable properties. Another relevant Armstrongian suggestion is that imperfect resemblance is a matter of different combinations of coinstantiated universals (1989: 103-107). To use a simple analogy, suppose that "pure" colors are universals, and that an orange block imperfectly resembles a red block with respect to color. On this Armstrongian proposal, the explanation of their imperfect similarity is that they both instantiate the red universal, but that the orange block also instantiates a yellow universal which differentiates them. Analogously, two glasses being imperfectly similar in terms of their fragility could be explained by one of them instantiating some other property that strengthens it. A disjunction of conjunctions of imperfectly similar universals could arguably be considered a determinable, and each of its disjuncts its determinates.

However, Armstrong argues that disjunctions of universals are not themselves universals (*ibid*.: 82-83). He claims that if two objects both have 'M or C' because one has mass M and the other has charge C, they do not really share anything identical, as must be the case if they share the same universal. Armstrong also argues against disjunctive universals on the grounds that such universals would make no difference to an object's causal powers. If an object has a certain causal power in virtue of having charge C, the disjunctive property 'M or C' "adds nothing to its power" (*ibid*.: 83). So, even if Vetter could say that a determinable potentiality is a disjunction of determinate universals, there is reason not to regard such potentialities as universals. Moreover, even if a less austere Immanent Realist than Armstrong were to allow that determinable properties are real, albeit non-fundamental universals, they would still run afoul Vetter's claim that determinable potentialities are more fundamental than their determinate degrees.

Another possibility is that determinable potentialities are higher-order universals. Consider a color analogy again. If a red block and a blue block have different color universals, perhaps 'having a color' is a higher-order universal that the blocks have in virtue of having some color or other. However, if the expression "in virtue of" indicates a grounding relation, then the determinable

property is grounded by its determinates. Determinable potentialities would be less fundamental than their determinates, which runs counter to Vetter's view. A related suggestion is that a determinable potentiality like breakability is a higher-order universal instantiated not by particular objects, but by other universals which are determinate degrees of breakability. This suggestion amounts to positing a hierarchy of universals, the higher-level ones being universals had by lower-level universals. Higher-order determinable universals are not instantiated by particulars directly, but indirectly, in virtue of their lower-order determinate universals being instantiated. The red block doesn't instantiate 'being a color,' but it instantiates redness, which is a color.

But would this allow potentialities to be located in their instances? Suppose that 'being a color' is a higher-order universal had by red and blue, and 'being a color' is located where the red and blue universals are – such as in the red and blue blocks, among other places. If so, then the determinable 'being a color' universal in the red block must be identical to the determinable 'being a color' universal in the blue block. If it is not, then determinables are not located in their instances. It strikes me as strange to think that objects of every color of the rainbow are identical with respect to the universal 'being a color' being located in each one of them. Having something strike you as strange probably counts even less in a philosophical debate than giving an incredulous stare, but perhaps it is enough to motivate a reconsideration of the approach to properties known as "Platonic" or "Transcendent" Realism.[12]

According to Transcendent Realism, universals *can* exist without being present in objects. One motivation for this claim is the intuition that it is contingent which properties things happen to have – they could have had different properties. If there are any properties that could have been instantiated in this world but aren't, then they are uninstantiated universals, on this Platonic view. Another motivation for uninstantiated universals is our grasp of kinds of perfection that are never realized. For example, while no actual, concrete object is perfectly circular, we can understand the property of being perfectly circular as an uninstantiated universal.

If we allow that there are uninstantiated universals, clearly these universals are not located in any actual, concrete object. Therefore, they must be elsewhere – outside of our actual space-time. But if that's the case, then maybe *instantiated* universals are outside of our space-time as well. Some uninstantiated universals differ from instantiated universals only contingently: One happens to be instantiated, while the other could have been, but isn't. These two otherwise similar universals exist in two different ways, in two different

[12] For an articulation of such a view, see Tugby 2013 and Giannini and Tugby (2020).

realms – that one is multiply located in actual concrete objects, while the other is not located in the actual, concrete world at all. What happens when a previously uninstantiated universal comes to be instantiated? Does its other-worldly existence cease once it acquires a worldly existence? If that seems problematic, it motivates the Platonist to "put" *all* universals where the uninstantiated ones are – outside of actual space-time (Armstrong 1989: 76). By the same reasoning, if determinable potentialities must be located somewhere other than their instantiations, then there is no bar to putting other universals there as well.

One advantage of Platonic Realism for Vetter is that some versions of the view already countenance the idea that universals are had in degrees, and so they can accommodate Vetter's view that potentialities comes in degrees. This is because one of the motivations for Platonism, as discussed above, is to explain the existence of an ideal that is never perfectly instantiated, like the perfect circle. If such an ideal property is never fully instantiated, that must mean that it is instantiated to a lesser degree. Coming in degrees is explained in terms of degrees of instantiation. The circle I draw with a compass instantiates circularity to a higher degree than the one I draw freehand. By the same token, Vetter could say that a wine glass instantiates breakability to a higher degree than a crowbar. How instantiation can be a matter of degree is a question I will not tackle here, but it is a question some Platonists already countenance, and not a new problem for Vetter (see Lloyd 1990).

This Platonic view is not flagrantly incompatible with Vetter's project. As higher-order universals, potentialities would be ultimately grounded in the ways that actual, concrete object are. However, they would not *be* ways that actual, concrete objects are. (Just as the red block doesn't have the property 'being a color,' a fragile glass doesn't have the property 'being a degree of breakability'.) Furthermore, it is difficult to see how standing in a relation to a universal that is located elsewhere could make a particular object powerful. Moreover, if we were to go for a hybrid Platonic view in which the determinate degree of potentiality were located in the object but the determinable potentiality were located elsewhere, then it would appear that the determinate potentiality that is intrinsic to the object is a more plausible source of the object's capabilities. This would make determinate degrees of potentiality better candidates for the source of modality than determinable potentialities – counter to Vetter's claim that determinable potentialities are more fundamental.

4. *The Trouble with Determinables*

Two constraints to finding an account of properties that is suitable for Vetter's project have emerged: 1. Conditions on property identity which employ

modal notions are problematic for Vetter's goal of grounding all of modality with potentialities; and 2. Theories of properties according to which properties are grounded in something else are inconsistent with Vetter's claim that nothing is more metaphysically fundamental than potentiality. In retrospect, the fact that the properties are supposed to be powerful plays a relatively minor role. The crux of the difficulty, from my point of view, is making sense of the idea that these relatively fundamental properties are determinables. Other philosophers argue against fundamental determinables on the grounds that determinables are asymmetrically necessitated or fixed by their determinates (Armstrong 2010: 50). Instantiation of a determinate, such as scarlet, entails that a determinable (red) is instantiated, but instantiation of redness does not entail that scarlet is instantiated. This asymmetry has implications for grounding relations. Since determinables do not determine their determinates, determinables cannot fully ground their determinates. Being red cannot fully ground being scarlet. Furthermore, as noted earlier (Lewis 1983), determinables are arguably less natural than their determinates, and this is reason to think that they are less fundamental. (I will consider arguments in favor of fundamental determinables below, in the context of assessing my positive proposal.)

The metaphysics of determinables remains unclear for Vetter. If they are sets, their status as fundamental is in doubt, for reasons discussed above.[13] They can't be universals located in their instances because then they would have to be identical across instances. They could be Platonic entities which objects instantiated to a greater or lesser degree; however, this would place the source of causal power outside of particular powerful objects.

An option we have yet to consider is Jessica Wilson's "subset view" according to which instances of determinable properties "have a proper subset of the powers of their associated determinate instances" (1999: 48)[14] This would solve some of the puzzles raised above, such as those regarding location, since a subset of properties is (at least partially) located where the set of properties is. Furthermore, since a set can be a subset of a great variety of different sets, including various ordered sets, we can make sense of the idea that potentialities come in degrees. However, it is not clear whether a view on which properties are sets of powers would work for Vetter's account of potentialities, for it would entail that potentialities are sets of powers. If the powers that are members of the sets are themselves potentialities, then potentialities are sets of potentialities *ad infinitum*. If these powers are *not* potentialities, then they

[13] For a version of Nominalism specifically developed to account for determinables, see Denby 2001.

[14] See also Shoemaker, Sydney, 2001, "Realization and Mental Causation", in Proceedings of the 20th World Congress in Philosophy, Cambridge: Philosophy Documentation Center, 23-33.

seem to be modal properties that are not grounded in potentialities, and thus Vetter's potentialities would fail to ground all of modality.

5. *Proposal: fundamental potentialities as necessarily manifesting*

I have been making the case that familiar approaches to properties are in tension with claims that Vetter makes about potentialities. Assuming that Vetter's project is worthwhile, this motivates making some adjustments, while attempting to preserve as much of her view as possible. One adjustment is to find or devise a new approach to properties (or lack-there-of) which would be more compatible with potentialities. Another way, the one I will explore here, is to revise some of Vetter's claims about potentiality so as to render them compatible with more familiar approaches to properties.

My proposal is to take up an idea that is suggested in *Potentiality*. On Vetter's view, some potentialities such as charge are nomological potentialities that are always possessed to the maximal degree. It is impossible for charge fail to manifest conformity with Coulomb's Law. She considers the option "to identify a nomological disposition such as electric charge with a determinate of the potentiality in question, namely its maximal determinate" (286). However, she rejects this option because it relies on what she regards as an unjustified assumption that determinate properties are more natural than their determinables (287). However, perhaps making this assumption would allow Vetter to solve some of the problems I have raised.

On this proposal, necessarily manifesting potentialities are the fundamental ground for determinable potentialities. Suppose that mass and charge are fundamental potentialities.[15] These potentialities necessarily manifest – they are always warping space-time, emanating an electrical field, or what have you. Fundamental potentialities are cumulative, so that when the same potentialities are instantiated in conjunction, their effect is multiplied. Consider the smallest unit of mass as an instantiation of a fundamental potentiality to warp space-time a tiny bit. The mass of composite entities is determined by the masses of their constituents, as is their potentialities to significantly warp space-time. Large quantities of mass are the bundling of small, and ultimately fundamental quantities of mass. Likewise, a large electrical charge is the additive effect of the accumulation of entities with unit electrical charge. (Whether or not this picture is empirically accurate is another question which I will consider below.)

The point here is that it is compatible with Immanent Realism, and perhaps Trope Nominalism as well. If every instance of unit charge is identical in quan-

[15] Vetter entertains this supposition as well (50).

tity, then the fundamental charge universal could be wholly located at each of its instances. (Alternatively, each instance could instantiate perfectly similar fundamental charge tropes.) Furthermore, this proposal accommodates Vetter's view that potentialities comes in degrees. There is a sense in which fundamental potentialities are at the extreme positive end of the spectrum in that they are necessarily manifesting. However, a single instantiation of a fundamental potentiality would be on the low end of the spectrum: Having only one unit of charge is the least charged anything could be. Higher degrees of charge are accounted for by multiple instantiations of unit charge.

On the view I am exploring, Vetter's determinable potentialities are not fundamental, but instead they are grounded by minimal, constantly manifesting potentialities. Much of the rest of Vetter's framework can be maintained. Nomological potentialities ultimately ground non-fundamental potentialities, including joint, extrinsic, and iterated potentialities. If determinable potentialities can ground all of modality as Vetter claims, then so could determinable potentialities grounded in determinate nomological potentialities. For example, if the mere possibility of someone being electrically shocked can be grounded by determinable potentialities, then it can be grounded in determinable potentialities which are in turn (partially) grounded by necessarily manifesting charge.

6. *Objections*

6.1. The fundamentality of the determinable

On the proposal sketched above, the potentialities that are possessed to the maximal degree are determinate, and they are more fundamental than determinable potentialities. Vetter holds that determinable potentialities are more fundamental than their determinate degrees. However, I have not found an argument for this claim in *Potentiality*.

Perhaps Vetter could avail herself of Jessica Wilson's arguments for fundamental determinables (Wilson 2012). Vetter writes

> Wilson (2012) uses similar considerations to argue that determinable properties may be just as fundamental as their determinates. I sympathize with Wilson's argument, and my argument is partly inspired by hers. However, I think I can make an even stronger case concerning the specific single-track and the general multi-track dispositions than Wilson can concerning determinates and determinables: the general dispositions are not only equally fundamental as the specific ones, they are more fundamental. (57)

I will examine Vetter's "stronger case" for fundamental multi-track dispositions next. But concerning Wilson's argument for fundamental determinables, Vetter acknowledges it only shows that some determinables are as fundamental *as their determinates*, and that they are *part of* the fundamental base. On Wilson's view, some determinates are fundamental too. Wilson writes "it is not plausible […] that [the fundamental] base will contain only determinable properties" (2012: 13). According to Wilson, determinate properties are needed to ground facts concerning how a given determinable instance is determined. Vetter's claims that potentialities are determinables and that "it's potentialities all the way down" entail the view that Wilson calls implausible.

In *Potentiality*, Vetter targets Alexander Bird's view that only maximally specific, "pure" dispositions are fundamental. According to Bird

> all impure dispositions are non-fundamental. Fundamental properties cannot be impure dispositions, since such dispositions are really conjunctions of pure dispositions, in which case it would be the conjuncts that are closer to being fundamental. (Bird 2007: 22)

Bird's view is that pure, single-track dispositions – dispositions with exactly one stimulus and one manifestation – are more fundamental than impure multi-track dispositions – dispositions that have multiple stimulus and manifestation pairs. In the context of arguing against the standard conception of dispositions, Vetter presents an argument against Bird's view that single-track dispositions are more fundamental than multi-track dispositions:

> 1. Scientific laws mention natural properties of varying degrees of fundamentality.
> 2. The relative fundamentality of laws correlates with the relative fundamentality of the properties that they reference.
> 3. The general law describing charge in functional terms (Coulomb's Law) is more fundamental than a law relating a specific degree of charge to other variables, because
> a. It holds a more central place in scientific practice;
> b. It is more explanatory;
> c. It is more parsimonious.
> 4. Coulomb's law references multi-track charge, not any specific single-track charge.
> 5. Therefore, multi-track charge is more fundamental than single-track charge. (50-58)

This lesson about charge extends to other nomonological potentialities, and to potentiality in general.

Since Vetter argues from her claim that ordinary dispositions are massively multi-track to her view that fundamental potentialities are determinables, I believe that she would endorse an argument for her view that determinable potentialities are more fundamental than their determinates along the same lines.[16] If this argument were successful, it would count against my proposal that determinable charge (having some degree of charge or other) is a conjunction of specific determinate charges, and the determinate charges are more fundamental. The argument against my proposal mirrors Vetter's argument against Bird's view, but for the last premise and the conclusion:

4*. Coulomb's law references determinable charge, not any specific determinate charge.
5*. Therefore, determinable charge is more fundamental than determinate charge.

One way to challenge this argument is to deny premise (4*). One can distinguish between mentioning properties and quantifying over them without mentioning them. A different way to interpret Coulomb's Law is that it is not referencing determinables at all, but is instead referencing the determinates by quantifying over them. The same goes for other general scientific laws.

But suppose that Vetter is right that general laws such as Coulomb's law reference determinable or multi-track potentialities, and these laws are more fundamental than their specific instances. Even so, I wonder what notion of relative fundamentality is at work. Are the relative-fundamentality relations between laws isomorphic to the relative-fundamentality relations between properties? Maybe what makes one law more fundamental than another is something different than what makes one property more fundamental than another. For example, relative fundamentality of laws might have more to do with the practice of science and less to do with metaphysics. Vetter appeals to the relative fundamentality between different branches of science – physics and chemistry, for example (56). Granted, physics is more fundamental than chemistry and that is some reason to think that properties mentioned in physics are more fundamental than chemical properties. But the laws considered above (Coulomb's Law and a specific instance of it) are both laws of physics. Maybe the more general law is more useful in science, and the

[16] Vetter clarifies the distinction between determinable and multi-track dispositions as follows: "the relation between the multi-track disposition and its many 'tracks' is not the same as that between a determinable (such as charge) and its determinates (such as electric charge, electric charge). Having a determinable property entails having one of its determinates, to the exclusion of all others. Having the multi-track disposition electric charge, on the contrary, entails having all the corresponding single-track dispositions". (53)

specific one is derivable. But if one thinks, as Vetter does, that laws are derived from potentialities[17] then how fundamental a law is will depend on how fundamental the relevant potentialities are. To argue that one potentiality is more fundamental than another because it figures in a more fundamental law would be to argue in a circle.

If one regards laws as mind-independent features of the universe, then it is not obvious that these laws stand in relations of relative fundamentality that correspond to the role that representations of such laws play in scientific practice. If one is a constructivist about laws, then the way that we rank our constructed laws in terms of fundamentality might not tell us much about the properties we mention in those laws. Furthermore, Vetter seems to assume that science is unified, and this assumption has been challenged.[18] If different sciences do not constitute one unified science, then a property may be fundamental in one science, but have no relative fundamentality relation to a property mentioned in another science. What's more, some sciences don't explain in terms of laws, but instead rely on context-based mechanistic explanations.[19] In such cases, there are no laws to provide a basis for relative fundamentality judgments.

Vetter claims that one reason to think that general laws are more fundamental than specific laws is that general laws are more explanatory than specific laws. However, it is not clear what these judgments of relative explanatory power are based on. If explanatory power rests on pragmatic considerations, then it is an unreliable guide to relative fundamentality. For example, consider Putnam's explanation of why a square peg won't fit into a round whole (Putnam 1975). In many contexts, citing the peg's shape is more explanatory than a detailed explanation of the relationships among the micro-constituents, but this is consistent with molecular constitution being more fundamental than macroscopic shape.

Furthermore, it is not clear what to make of the considerations of parsimony. In some sense, it is much more parsimonious to make a universal generalization than it is to describe every positive instance of that generalization. Nevertheless, many philosophers think that some universal generalizations are grounded by their instances, and Vetter concedes that grounds are more fundamental than that which they ground (23; see also Bennet 2017: 208). Laws that are more general and inclusive may be simpler and more useful than de-

[17] See also Bird 2007; Cartwright 2009; McKitrick 2005.

[18] For arguments that science is disunified see Dupré 1993; Galison and Stump 1996; Cartwright 1999.

[19] For example, see Lombrozo 2010; Woodward 2003.

scriptions of specific cases. But if laws are just certain kinds of generalizations, then they are less fundamental than what they are generalizing over. Hence I question premise 3. We should be wary of equating metaphysical fundamentality with being fundamental to scientific practice.

6.2. Empirical considerations

While I do not think that Vetter has a good argument against fundamental determinate potentialities, I have some remaining worries. In general, I worry that my proposal is incompatible with certain empirical possibilities. This would be a problem for Vetter adopting my proposal, since she wants to use potentialities to give an account of possibility. The proposal might have the consequence that something that seems possible turns out to be impossible. That might be a bullet worth biting, but if any of these possibilities turn out to be actual, the view would be falsified. One such possibility is that mass is fundamental, but there is no unit mass. Instead, there are multiple kinds of simple entities that have different masses. So, the mass of a more massive simple entity cannot be reduced to smaller units of mass. A related possibility is that some quantities, such as mass, are continuous. One could say that there are infinitely many fundamental mass properties of different quantities – different mass universals. These would be functionally similar potentialities differing only in degree, and so it would be impossible to explain, in terms of universals, what they have in common.

Perhaps these scenarios are possible or even actual, and my proposal to save Vetter's view does not work. On the other hand, Vetter wants to ground modality in the ways that actual objects are, and potentially are. So, I think that it is inevitable that her view has some empirical implications, and so it can be falsified by empirical facts. It follows that empirical facts constrain what is possible. This brings Vetter's view closer to that of Dispositional Essentialists such as Brian Ellis (2001: 203-257) and Alexander Bird (2007: 43-59) who dissolve the distinction between physical and metaphysical possibility.[20]

7. *Conclusion*

In this chapter, I tried to show that it is difficult to find a metaphysics of properties that provides the kinds of properties that Vetter calls "potentialities". In the end, I think that Vetter should give up on the idea that fundamental potentialities are determinable, but instead embrace the idea that she

[20] Vetter notes that rejecting the view that the laws of nature of metaphysically contingent is "the natural move for the potentiality view" but she argues that she is not committed to doing so (282).

already entertains – that fundamental (nomological) potentialities are necessarily possessed to the maximal degree. The maximal degree is a determinate degree, not a determinable. The determinable property is grounded by cumulative instantiations of determinate properties. However, if this metaphysics clashes with our best scientific evidence, more radical rethinking of *Potentialities* may be required.[21]

Jennifer McKitrick
jmckitrick2@unl.edu
University of Nebraska-Lincoln

References

Armstrong, D.M., 2010, *Sketch for a Systematic Metaphysics*, Oxford University Press, Oxford.

Armstrong, D.M., 1997, A *World of States of Affairs*, Cambridge University Press, Cambridge.

Armstrong, D.M., 1989, *Universals: An Opinionated Introduction*, Westview Press, Boulder CO .

Armstrong, D.M., 1978 *Universals and Scientific Realism*, Cambridge University Press, Cambridge.

Bennett, K., 2017, *Making Things Up*, Oxford University Press, Oxford.

Bird, A., 2007, *Nature's Metaphysics: Laws and Properties*, Oxford University Press, Oxford.

Bird, A.. "Dispositions and Antidotes", in *The Philosophical Quarterly*, 48 (1998): 227-234.

Cartwright, N., 2009, "Causal Laws, Policy Predictions, and the Need for Genuine Powers", in T. Handfield (ed.), *Dispositions and Causes*, Oxford University Press, Oxford: 127-157.

Cartwright, N., 1999, *The Dappled World: A Study of the Boundaries of Science*, Cambridge University Press, Cambridge.

Contessa, G., 2013, "Dispositions and Interferences", in *Philosophical Studies*, 165, 2 (2013): 401-419.

Correia, F., and B. Schnieder, eds., 2012, *Metaphysical Grounding: Understanding the Structure of Reality*, Cambridge University Press, Cambridge.

Denby, D.A., 2001, "Determinable Nominalism", in *Philosophical Studies*, 102: 297-327.

Dupré, J., 1993, *The Disorder of Things. Metaphysical Foundations of the Disunity of*

[21] Thank you to two anonymous referees, Dominic Bailey, Nick Jones, Jessica Wilson, and those who attended my presentation of this material at The Ranch Metaphysics Workshop for their comments.

Science, Harvard University Press, Cambridge MA.

Eddon, M., 2013, "Quantitative Properties", in *Philosophy Compass* (2013) 8: 633-645.

Ellis, B., 2001, *Scientific Essentialism*. New York: Cambridge University Press: 203-257.

Galison, P. and D. Stump, eds., 1996, *The Disunity of Science. Boundaries, Contexts and Power*, Stanford University Press, Stanford CA.

Giannini and Tugby, D., 2020, "Potentiality: actualism minus naturalism equals platonism", in *Philosopihical inquiries*, 8, 1: 117-140.

Lewis, D., 1997, "Finkish Dispositions", in *Philosophical Quarterly*, 47 (1997): 143-158.

Lewis, D., 1983, "New Work for a Theory of Universals", in *Australasian Journal of Philosophy*, 61: 343-377.

Lloyd, A.C., 1990, *The Anatomy of Neoplatonism*, Clarendon Press, Oxford.

Lombrozo, T., 2010, "Causal-Explanatory Pluralism: How Intentions, Functions, and Mechanisms Influence Causal Ascriptions", in *Cognitive Psychology*, 61: 303-332

Manley, D., and Ryan Wasserman, 2007, "A Gradable Approach to Dispositions", in *The Philosophical Quarterly*, 57 (2007): 68-75.

Martin, C.B., 1977, "Dispositions and Conditionals", in *Philosophical Quarterly*, 44 (1994): 1-8.

McKitrick, J., 2018, *Dispositional Pluralism*, Oxford University Press, Oxford.

McKitrick, J., 2019, "Review of Barbara Vetter's Potentiality: From Dispositions to Modality", in *Erkenntnis*, 84: 1179-1182.

McKitrick, J., 2005, "Introduction to Dispositions and Laws of Nature", in *Synthese*, 144: 305-308.

Mumford, S. and Anjum, R.L., 2011, *Getting Causes From Powers*, Oxford University Press, Oxford.

Putnam, H., 1975, "Philosophy and our mental life", in *Mind, Language, and Reality*, Cambridge University Press, Cambridge: 291-303.

Rodriguez-Pereyra, G., 2016, "Nominalism in Metaphysics", in E.N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (Winter 2016 Edition), URL = <https://plato.stanford.edu/archives/win2016/entries/nominalism-metaphysics/>.

Shoemaker, S., 2001, "Realization and Mental Causation", in *Proceedings of the 20th World Congress in Philosophy*, Philosophy Documentation Center, Cambridge: 23-33.

Smith, A.D., 1977, "Dispositional Properties", in *Mind*, 86 (1977): 439-445.

Tugby, M., 2013, "Platonic Dispositionalism", in *Mind*, 122(486): 451-480

Vetter, B., 2015, *Potentiality. From Dispositions to Modality*, Oxford University Press, Oxford.

Whittle, A., 2009, "Causal Nominalism", in T. Handfield (ed.) *Dispositions and Causes*, Oxford University Press, Oxford.

Wilson, J. M., 2012, "Fundamental Determinables" *Philosophers' Imprint*, 12 (2012), 1-15.

Wilson, J. M., 1999, "How Superduper Does a Physicalist Supervenience Need to Be?" *The Philosophical Quarterly*, 49(194): 33-52.

Woodward, J., 2003, *Making Things Happen: A Theory of Causal Explanation*, Oxford University Press, Oxford.

Potential problems? Some issues with Vetter's potentiality account of modality

Nathan Wildman

Abstract: As Vetter says, we are at the "beginning of the debate, not the end" (2015: 300) when it comes to evaluating her potentiality-based account of metaphysical modality. This paper contributes to this developing debate by highlighting three problems for Vetter's account. Specifically, I begin (§1) by articulating some relevant details of Vetter's potentiality-based view. This leads to the first issue (§2), concerning unclarity in the idea of degrees of potentiality. Similarly, the second issue (§3) raises trouble for Vetter's proposed individuation conditions for potentialities. Finally, the third issue (§4) is about apparently unmanifestable intrinsic potentialities, and suggests that there might be some deeper problems with anchoring metaphysical possibilities in concrete objects. More generally, though the issues detailed here are problematic, I do not take them to be fatal. However, they do show that, at minimum, further clarification of Vetter's potentiality view is required.

Keywords: possibility; metaphysical modality; potentiality; dispositionalism; Vetter, Barbara.

According to dispositionalism, starting from the broadly modal notion of a disposition, we can offer an analysis of metaphysical modality – i.e., of metaphysical possibility and necessity. And, as a knock-on consequence, one can also account for the rest of the 'modal package' – i.e., the counterfactual conditional, essentiality, laws of nature, etc.

In many ways, dispositionalism is an attractive approach to modality. For one, it is ideologically parsimonious: it has only one primitive – *dispositionality* – to which everything else reduces (or can at least be defined in terms of). For another, it promises an account of metaphysical modality in terms of actual, concrete objects and their properties. Consequently, dispositionalists do not need to postulate non-actual entities (e.g. the various denizens of Lewisian possible worlds) to serve as the anchors for modality. Finally, and relatedly, by anchoring modality in the dispositions of ordinary, actual objects, dispositionalism offers an extremely plausible epistemology of modality. Specifically, we can and do engage in empirical investigation to determine the dispositions of every day,

ISSN (print) 2281-8618-ETS DOI: 10.4454/philinq.v8i1.280

and given dispositionalism, the epistemology of metaphysical modality is just a generalization of this process. This is a particularly appealing result because many of the competing accounts of modality – e.g. Lewisian realism and Finean essentialism – make the epistemology of modality extremely mysterious.

For these (and other) reasons, a number of philosophers have recently begun developing versions of dispositionalism. This includes Bird (2007), Pruss (2002), Borghini and Williams (2008), Jacobs (2010), and Anjum and Mumford (2018).

But however appealing dispositionalism is, it also faces a number of difficulties. Chief among them is providing suitable analyses of the core modal notions of possibility and necessity in terms of dispositions. For example, it is *prima facie* plausible that, if an object *a* has a disposition to *M*, then, possibly, *a* is *M*.[1] However, it is not at all clear how to extend the story so as to capture every possibility. For example, it is possible that the Golden Gate Bridge breaks, despite the fact that the Bridge intuitively lacks the disposition to break. In this way, the challenge for the dispositionalist is to provide a "full-fledged account of modality that is true to the spirit of dispositionalism without flying in the face of our most central pre-theoretical beliefs about what is possible" (Contessa 2016: 1238).

Barbara Vetter (2015) offers an ingenious new version of dispositionalism which seems to solve the challenge. Specifically, Vetter suggests that we replace the usual conception of disposition with that of *potentiality,* which includes, but 'extend[s] beyond' dispositions and abilities (2015: 142). Appealing to potentialities allows Vetter to offer a definition of (metaphysical) possibility in terms of potentialities:

POSSIBILITY It is possible that P iffdf something has, had, or will have an iterated potentiality for it to be the case that P (2015: 199)

From this, definitions for various other modal notions can then be constructed. For example, Vetter defines necessity as:

NECESSITY It is necessary that P iffdf nothing has, had, or will have a potentiality that not-P (2015: 203)

In this way, Vetter's potentiality-based account promises to be an extremely valuable contribution to the larger project of making sense of (metaphysical) modality; it has all the benefits of dispositionalism while apparently circumventing its biggest problems.

[1] In fact, even this first step is questionable, as there may be impossible dispositions. See e.g. Jenkins and Nolan (2012) and Vetter (2015: 250-257; 2016) for further discussion.

Yet we are, as Vetter says, at the "beginning of the debate, not the end" when it comes to assessing her potentiality-based view (2015: 300). While Vetter offers us an innovative and nuanced dispositionalist account of modality, we do not yet have a complete and perfectly clear picture about whether the potentiality view is a better option than the competition. For example, it is not obvious how the potentiality view fares when we compare it to an essence-first approach, as developed by e.g. Fine (1994a, 1994b, 1995, 2000) and Correia (2006, 2012), or with Lange's (2009) subjunctive fact-based story.[2] Nor have all of the potentiality view's various wrinkles been ironed out yet. For example, is Vetter's view as ideologically simple a story as it first appears? And can we really get satisfactory definitions of other modal notions out of potentiality?

Furthering this debate involves determining whether the benefits of Vetter's potentiality account are worth the "costs" – i.e., whether the theoretical gains we make from adopting the position sufficiently counter-balance the problematic or counter-intuitive results that the theory entails.[3]

This paper is a contribution to this debate.[4] My aim is to highlight some potential difficulties for Vetter's account, thereby indicating some potential costs that would be incurred, were one to adopt Vetter's position. Specifically, I here identify three issues. While these issues are problematic (and substantive), I do not take them to be *fatal*; that is, I don't think they prove that the potentiality-based view is doomed. However, they do show that, at minimum, further clarification of the view is required.

The plan is as follows. I begin (§1) by articulating some relevant details of Vetter's potentialist view. This leads to the first issue (§2), which concerns a lack of clarity regarding Vetter's idea that potentialities admit of degrees. A natural way to understand this degree-talk is in terms of proportions of possible worlds; however, this leads to counter-intuitive results. In the end, exactly how to understand degrees remains mysterious. The second issue (§3) concerns individuation conditions for potentialities. Here, I argue that potentiality individuation is more complex than Vetter posits, which suggests that the position is not as ideologically parsimonious as it might first appear. Finally,

[2] Notably, Vetter (2015: §5.6) discusses how her conception of potentiality relates to Fine's view about essence, but she does not explicitly compare the two overall accounts of modality.

[3] The methodology here is something that was exemplified by Lewis (1986), but is neatly summarized by Sider: "Competing positions are treated as tentative hypotheses about the world, and are assessed by a loose battery of criteria for theory choice. […] Theoretical insight, considerations of simplicity, integration with other domains (for instance, science, logic, and philosophy of language), and so on, play important roles" (2009: 385).

[4] For other objections to Vetter, see e.g. Schrenk (2015), Contessa (2016), Leech (2017), and McKitrick (2019).

before concluding, I raise (§4) a third issue, about apparently unmanifestable intrinsic potentialities. This suggests that there might be some problems with the general idea that we should anchor possibilities to objects.

1. *Background: Vetter on potentialities*

To understand Vetter's picture of the nature of potentialities, it is helpful to start by considering the standard view of dispositions. The standard, conditional analysis of (single-track) dispositions defines dispositions via counterfactual conditionals. So, "o is disposed to M" is, on this view, defined as something like, "If it were the case that S, then o would M". In this way, the standard analysis individuates dispositions by both their stimulus conditions – the "S" in the counterfactual's antecedent – and their manifestation – the "M" in the consequent.

Vetter rejects this conditional analysis. The problem, according to Vetter, is that, given the massive qualitative and quantitative diversity of suitable conditions, it is not clear how to adequately specify the stimulus conditions under which all and only the relevantly disposed things would fulfil the manifestation condition. After exploring a number of ways to try and do so, Vetter argues that they all fail. There is, per Vetter, no way to identify the right stimulus conditions.

Building off of this, Vetter suggests abandoning the appeal to stimulus *and* manifestation conditions. Her alternative starts from the idea that we approach dispositions in terms of the *potentiality* to manifest a relevant condition (2015: 65). Potentialities are, according to Vetter, individuated purely in terms of their manifestations. And for an object *x* to have the potential to F just means that, possibly, *x* Fs. For example, an object *x* is breakable iff *x* has the potential to break – and to have the potential to break just means that, possibly, *x* breaks. And the circumstances that might bring about *x*'s breaking are irrelevant – all that matters is that *x can* break.

Of course, many things have the potential to break – both a delicate vase and the Golden Gate Bridge *can* break, if subject to enough force.[5] But what differentiates the two is the *degree* to which they have this potential: while both have the potential to break (i.e., they both can, in the metaphysical sense of 'can', break), the vase is breakable to a greater degree than the Bridge. These degrees range from the minimal degree, which is simply possibly manifesting the relevant condition, to the maximal degree, which is having no potential whatsoever to not manifest the relevant condition (Vetter 2015: 90).

[5] Arguably, *every* concrete object can, in the metaphysical sense of 'can', be broken.

This difference in degree also helps explain why it is that only the vase has the disposition of *fragility*. This is because our everyday dispositional attributions are (mostly) contextual:

> For some disposition terms, such as 'fragile', a given context imposes a threshold: how fragile an object has to be in order to count as fragile simpliciter. An object x counts as fragile in a context C iff x is above that threshold. For other disposition terms, such as 'breakable', any positive proportion is suitable, and no contextual threshold is required. (Vetter 2015: 78)

So, what demarcates something as fragile in an everyday context is that (i) it has the potential to break, and (ii) it does so sufficiently easily – where this 'ease' is determined by some contextually variant standard. And what guarantees the satisfaction of both conditions is that the relevant object has the potential to a sufficiently high degree.

Allowing potentialities that admit of degrees allows Vetter to directly address problems that plague more traditional dispositionalist accounts of modality. Most importantly, it no longer need be the case that, for an object x to possibly be F, x must be *disposed* to F; instead, x need only *potentially* be F. So Vetter can (rightly) say that "The Golden Gate Bridge is fragile" is false – the Bridge does not have the potential to break to a sufficiently high degree to satisfy the contextually determined threshold such as to make the Bridge fragile. And she can do so while maintaining that "Possibly, the Bridge breaks" is true, in virtue of the fact that the Bridge has the potential to break to some degree.

Similarly, consider cases where, for example, a fragile vase doesn't break when struck gently. Vetter can (rightly) hold that the vase is still fragile, because the vase still has the potential to a sufficiently high degree to satisfy the contextually determined threshold for fragility. All that has happened is that the amount of force applied in this case did not suffice to cause the vase to manifest this potential.

What we have then is something like the following picture. The fundamental elements are potentialities, which are (i) properties possessed by individuals, (ii) individuated by manifestation conditions, (iii) such that they admit of degrees, and (iv) are closely tied to possibility. In particular, *x*'s having the potential to *F* entails that possibly, *x F*s. Dispositions, meanwhile, can be thought of as a sub-type of potentialities. Specifically, to have a disposition is to have the potential to manifest a certain condition to a certain degree, a degree which is often contextually determined. In this way, while potentialities are individuated purely by manifestation condition, dispositions are

individuated by manifestation condition and degree.[6]

This is an extremely quick sketch of Vetter's account of potentialities and dispositions. It glosses over many details, and only gives a rough approximation of several others. But it suffices to give us a grip on the foundations of her potentiality-based account. More importantly, it highlights certain points about the view which are directly relevant to the problems I will raise in the next section.

2. *Degrees of potentiality?*

A central plank in Vetter's account is the idea that potentialities come in degrees. And the first issue that I would like to raise concerns this notion of degrees.[7] Specifically, it is not at all clear to me how exactly we ought to understand this talk of degrees.

For example, suppose I am sitting at my desk with a ceramic coffee mug at one elbow and a glass beer mug at the other. Which of the two mugs is *more easily* broken – i.e., which has the potential to break to a great degree?

One option is to consider proportions of worlds. Specifically, on a "proportional conception of degrees, *x* is more [breakable] than *y* just in case *x* breaks in more of the relevant worlds than *y*" (2015: 73). Of course, this just leads to the question of which are the 'relevant' worlds. Thankfully, Vetter suggests the following explication:

> *x* is more [breakable] than *y* just in case the proportion of worlds where *x* has its relevant intrinsic features and breaks is greater than the proportion of worlds in which *y* has its relevant intrinsic features and breaks. (Vetter 2015: 78)

Extending this idea into a general principle gives us:

> PROPORTION *x* has potentiality *P* to a greater degree than *y* iff the proportion of worlds where *x* has its relevant intrinsic features and *P*s is greater than the proportion of worlds where *y* has its relevant intrinsic features and *P*s

This gives us a way to spell out talking of degrees of potentiality. Suppose there are 100 worlds where beer mug *b* has intrinsic features *I*, and, in 35 of

[6] When discussing her view, Vetter often talks as if we can individuate dispositions solely by their manifestations, but she is clear that this is a 'simplification' (2015: 96), and that we must appeal to degrees too.

[7] As will become clear shortly, this worry is related to, but distinct from, concerns about measuring proportions involving potential infinities; see (Manley and Wasserman 2008: 79-81) and Vetter (2014: 141-143; 2015: 77-78) for discussion.

these worlds, *b* breaks. Meanwhile, suppose that there are 100 words where coffee mug *c* has intrinsic features *I'*, and, in 25 of these worlds, *c* breaks.[8] Because the proportion of *b*-*I* worlds where *b* breaks is larger than the proportion of *c*-*I'* worlds where *c* breaks, *b* has the potential to break to a higher degree than *c*.

Note that PROPORTION can be read in two ways. The *reductionist* reading treats the principle as providing a reductive definition of 'more easily', and hence a reductive account of degrees of potentiality. For obvious reasons, Vetter rejects this reductionist reading: she wants an account of modality in terms of potentials, so if she reduces potentials to some other (broadly) modal notion, then she undermines her own would-be foundation.

Instead, Vetter suggests that we take the ordering of potentiality degrees as primitive, and adopt the *realist* reading, according to which the principle merely specifies a "formal model and rough approximation" of degrees of potentiality (Vetter 2015: 78).

In what follows, I, following Vetter, will understand PROPORTION in the realist, rather than reductionist manner. That said, it is worth noting that, even if we accept the realist reading, because its main connective is an 'iff', if PROPORTION is true, then it must specify a logical equivalence between talk of degrees of potentiality on the one hand and proportions of (relevant) possible worlds on the other. So, even the realist who accepts PROPORTION must agree that, at minimum, you cannot have a difference in truth-value between (relevant) claims about world-proportions and claims about potentiality degrees.[9] (Obviously, a realist might reject PROPORTION entirely, but more on that option in a moment.)

The problem I'd like to highlight is that, even if we just read PROPORTION in this realist manner, we can generate counter-intuitive evaluations of potentiality degree-talk. To see the problem, it's helpful to consider the following (slightly idiosyncratic) science fiction example.

Take Asmodeus, a king cobra (Ophiophagus hannah), whose bite is extremely venomous to humans – it is capable of delivering enough neurotoxins to kill an Asian elephant, as well as 50 percent of the humans she bites. Meanwhile, Basil (short for Basilisk) is a *cyber king cobra* – part animal, part machine – whose

[8] Obviously, these numbers are massive simplifications.

[9] This entails that, strictly speaking, the realist cannot both accept PROPORTION and take it to be a mere "rough approximation" (Vetter 2015: 78). Rather, the realist must either take PROPORTION to be true, in which case it specifies the relevant logical equivalence, or to be false and at best a close approximation to the truth. Of course, opting for the latter leaves us with no answer to the overall question of how we should understand degrees of potentiality but also implies that there is a re-formulated principle – call it PROPORTION* – that is true. And one way to read the point of this section is as asking the would-be realist to spell out PROPORTION*.

animal parts were grown in a lab (and are derived from a real king cobra's) and whose high-tech nanomachines were developed by human scientists (perhaps working for the Tyrell or Wallace Corporations). Like Asmodeus, Basil's bite is extremely venomous to humans. However, he is not *quite* as venomous – perhaps because of the presence of the nanomachines in Basil's venom, his bite is only capable of killing 45 percent of the humans he bites. In this way, Basil's bite is, intuitively, less venomous to humans than Asmodeus' – in other words, she has the potential to a greater degree than he does.

When we consider the proportion of worlds where Asmodeus has the particular intrinsic properties she actually does and she manifests her venomous-to-humans potential, it will be suitably high. She is, after all, one of the most venomous snakes in the world! Of course, it won't be *that* high. And the biggest factor dragging the proportion down is that not all king cobra-populated worlds are also human-populated worlds. Consequently, there are a significant number of worlds where Asmodeus has her specific intrinsic profile and there are no humans, meaning she will not be able to manifest her potential.

Like with Asmodeus, when we consider the proportion of worlds where Basil has the particular intrinsic properties he actually does and he manifests his venomous-to-humans disposition, it will be suitably high. However, it won't be *that* high – he only kills 45% of the people he bites, after all. And it clearly should be lower than Asmodeus' proportion.

But there's a complication: Basil stands in an ontological dependence relation to humans. More specifically, assuming that Basil is essentially a cyber king cobra, by dent of his (essential) nanomachine parts, Basil and his ilk can only exist in worlds where the nanomachines exist. And, as the nanomachines are artefacts, they only exist in worlds where their creators – i.e., humans – do too. So, all cyber king cobra worlds are human worlds.

The upshot is that the biggest factor impacting Asmodeus' proportion is not present in the case of Basil. Consequently, Basil is effectively guaranteed to have a larger proportion than Asmodeus. Given PROPORTION, this entails that Basil has the relevant potential to a greater degree than Asmodeus. But this is exactly opposite of the intuitive outcome!

More generally, let P be a potential whose manifestation involves entities of kind K in some way, a an object that has P to some non-maximal and non-minimal degree, and b an object that (i) has P to a slightly lesser degree than a, and (ii) is such that it cannot exist without there being K's.[10] The proportion for a is defined as the number of worlds where a has its relevant intrinsic fea-

[10] The easiest way to guarantee satisfying clause (ii) is to make b generically ontologically dependent upon K's.

tures and manifests *P* out of the total number of worlds where *a* has its relevant intrinsic features. The former consists entirely of *a*-and-*K* words, though the latter includes both *a*-and-*K* and *a*-without-*K* worlds. In other words, the proportion is the following:

Number of worlds where *a* manifests *P* and *K*'s exist	:	Number of worlds where *a* manifests *P* and *K*'s exist + Number of worlds where *a* does not manifest *P* and *K*'s exist + Number of worlds where *a* does not manifest *P* and *K*'s do not exist

Given plausible assumptions about modal variation and plenitude, the largest number here by far will be that of the worlds where *a* does not manifest *P* and *K*'s do not exist.

Meanwhile, *b*'s proportion is defined as the number of *b*-manifests-*P*-worlds over the number of worlds where *b* does not manifest *P*. Here, because *b* cannot exist without there being Ks, both the former and latter will consist of *b*-and-*K* words. That is, this proportion is something like:

Number of worlds where *b* manifests *P* and *K*'s exist	:	Number of worlds where *a* manifests *P* and *K*'s exist + Number of worlds where *a* does not manifest *P* and *K*'s exist

Importantly, the denominator is obviously much smaller than in the previous case, since the number of worlds where *b* does not manifest and *K*'s do not exist is zero. Consequently, this proportion is guaranteed to be larger than the one for *a*. Given PROPORTION, it follows that *b* has *P* to a greater degree than *a*. But this contradicts the initial stipulation that *b* has *P* to a lesser degree than *a*.

The possibility of such cases strongly calls into question understanding talk of potentiality degrees in terms of talk of proportions of worlds.

Of course, a reply seems ready to hand: Vetter can say (as she does in reply to a different, but related objection) that this is just "another shortcoming of trying to account for such perfectly intelligible notions as that of a disposition in terms of possible worlds" (2015: 78). Instead, we should take the ordering of potentiality degrees as primitive (2015: 81), and leave aside any attempt to cash out degree talk in other terms.

This certainly circumvents the above problem: if we forgo trying to make sense of potentiality degrees in terms of world proportions – even in the

weak, realist manner – then we cannot generate the above problem.[11,12]

However, it also leaves us with no way to translate between potentiality-degree talk and world-talk. This makes degrees of potentiality mysterious: if we cannot cash degrees out in terms of proportions of worlds, how, exactly, should we understand them? We have no answer.

Of course, Vetter will likely respond that this talk of degrees is "perfectly intelligible" (2015: 78) on its own, without any such story linking it to things like proportions of worlds. Yet for those of us who struggle to understand potentiality degrees, this is cold comfort.[13]

Let us summarize. The idea that potentiality admits of degrees is central to Vetter's account; in particular, it plays a key role in ensuring that she has an extensionally adequate theory with regards to possibility claims. However, it is not clear how best to understand this talk of degrees. A natural way to do so is in terms of proportions of possible worlds, as in PROPORTION, which specifies a logical equivalence between certain degree claims and claims about proportions of worlds. The problem is that it is possible to generate counter-examples to this logical equivalence, as exemplified by the Asmodeus-Basil case. The most Vetter-friendly response to such counter-examples is to give up on PROPORTION and insist that (i) degrees of potentiality be taken as primitive, and (ii) claims about degrees are not logically equivalent to any claims about (proportions of) worlds. But this just makes potentiality degrees even more mysterious.

[11] In this way, one can read this point as simply strengthening Vetter's case for our being better off thinking of potentiality degrees as primitive.

[12] An alternative response, suggested by an anonymous referee, would be to insist that every entity can fail to co-exist with any other entity, either directly or via counter-parts. This could be motivated by appeal to a broadly Humean theory of recombination. Obviously, this would eliminate the difference between Asmodeus and Basil, since the case turns on the latter not being as modally "free" as the former. However, this does not seem like a suitable move for someone like Vetter, since it relies upon approaching modality via recombination, rather than via the potentialities of actual, concrete objects.

[13] Another potential response is to suggest that the counter-example case is built around the assumption that the relevant venomous potentiality is manifested in the killing of humans. But, as an anonymous referee has suggested, why understand the disposition in this way – couldn't the same disposition be manifested by killing some other animals? Two points in reply. First, while this might block the Asmodeus-Basil example, it is not obvious how to extend the response to block every version of the objection we can generate using the above schema. Second, as we will see in the next section, to avoid a different objection, Vetter must buy into fine-grained manifestation conditions. Yet once we start thinking that manifestations are fine-grained, it is not obvious how Vetter could block a version of the Asmodeus-Basil case that insisted upon using *venomous-to-humans* as the potential, rather than simply venomous. For at the fine-grained level, these two have different manifestations, and hence are different potentialities.

3. *Individuation of potentialities?*

The second issue I would like to raise concerns the individuation of potentialities. To get a grip on it, it is helpful to quickly talk through a different objection.

Consider coulrophobia (fear of clowns) and cynophobia (fear of dogs). Both seem to have the same manifestation: namely, being afraid. So, according to Vetter's account, they are the same potential. Further, if we stipulate that they are possessed to the same degree, it follows that these are the same disposition on Vetter's account. However, they are intuitively distinct. And what distinguishes them seems to be their particular stimuli – exposure to clowns and to dogs, respectively – which suggests that we need to return to something like the counterfactual conditional account.[14]

Vetter's response (2014: 149, 2015: 78fn14) is to say that the relevant manifestations are more complex than they first appear, often incorporating elements that look a lot like the stimuli. Specifically, coulrophobia's manifestation is *being afraid of clowns*, while cynophobia's manifestation is *being afraid of dogs*. As these manifestations are distinct, the account does not identify the two after all.

But consider *perishable* and *destructible*.[15] The former's manifestation condition is to *perish* – i.e., to go out of existence[16] – and the latter's manifestation is to *be destroyed* – i.e., to go out of existence. However, the former is nearly always attributed to entities that are (or were) alive, and the latter nearly always to (non-living) artefacts. This difference in application strongly suggests that the two are distinct: one is a potential only possessed by (formerly) living things, while the latter only by non-living entities. However, Vetter's account entails that they are identical, since the two have the same manifestation condition.

More generally, there seem to be some potentialities that have the same manifestation condition but are intuitively distinct because they apply to different kinds or sorts of entities. Vetter's account entails that the two are the same potential/disposition, which is an unpalatable result. This suggests that the individuation conditions for potentialities are more complex than Vetter suggests. Consequently, it would be good if Vetter could spell out what exactly these conditions are.

[14] Vetter (2014: 149) credits this objection to Alastair Wilson and an anonymous referee.

[15] Thanks to Stephan Leuenberger for suggesting this pair of dispositions.

[16] Etymologically, the English comes from the Latin present active infinitive of *pereo*, which is itself a combination of *per* ('through') and *eö* ('to go').

4. *Unmanifestable intrinsic potentialities?*

To build up to the third issue, it is helpful to spell out a few more details about potentialities. Specifically, Vetter thinks that there are a number of different types of potentialities. The simplest are *intrinsic potentialities* – i.e., potentialities that are intrinsic to the object that possesses them. Along with these intrinsic potentialities, there are *joint potentialities*, potentialities that two or more things jointly possess (e.g., Vetter and I possess the joint potentiality of our singing a duet).[17] These joint potentialities ground *extrinsic* potentialities,[18] which concern individuals external to the object that possesses the potentiality. For example, Vetter possesses the extrinsic potentiality that I sing a duet, and this extrinsic potential is grounded in our joint potential to sing a duet. Finally, there are *iterated potentialities*. Iterated potentialities are potentials to have certain potentials.[19] For example, Vetter has the iterated potentiality to have a daughter who has the potentiality to play the piano. And it is these iterated potentialities that directly bridge potentialities to possibilities, in POSSIBILITY.

The third issue concerns certain intrinsic potentialities like perishable, destructible, and mortal. Specifically, it does not seem possible for the objects that possess these potentialities to manifest them. But if it is not possible for the objects to manifest them, then it is not clear how the relevant story goes such that we can derive the appropriate possibility claims.[20]

Take my mortality. Plausibly, this disposition expresses an intrinsic potentiality of mine: namely, that I have the potential to die. According to Vetter's account, I have the potential to die iff I can manifest *M*, where *M* is mortality's characteristic manifestation condition. However, there is no obvious property *M* that I can manifest. Realizing my mortality involves my death, which (plausibly) involves my ceasing to be. But if I cease to be, then I am not around to be doing any manifesting. So I cannot in fact manifest the potential.[21]

[17] Joint potentialities are often grounded in the individual potentialities of their possessors, though Vetter gives no general story about the ground link between the two types due to a variety of issues. For further discussion see Vetter (2015: §4.3.4)

[18] For more on the grounding connections between joint and extrinsic potentialities, see Vetter (2015: §4.5).

[19] It is clear that at least some iterated potentialities are grounded in joint potentials; for example, my having the iterated potentiality to be such that I potentially am friend

[20] In this way the following objection is related to, though distinct from, problems that might emerge from apparently impossible dispositions as in Jenkins and Nolan (2012).

[21] Note that the problem I am raising here is a different (though related) problem to the one discussed by Cameron (2008) about all actual existing contingent beings not existing, and by Leech (2017) and Kimpton-Nye (2018) concerning potentialities for a thing to never have existed at all. Vetter (2015: 274) thinks the latter problem is particularly troublesome; my discussion here is an attempt to show that the "simple" problem is more problematic than she thinks.

More precisely, the first-order being constraint says that, necessarily, if an object has a property, then it exists. This principle is a generalization of the appealing and familiar idea that something has to be in order be a certain way, and has the air of something almost analytic; as Williamson puts it, "How could a thing be propertied were there no such thing to be propertied? How could one thing be related to another were there no such things to be related?" (2013: 148).

Suppose, for *reductio*, that F is the property I possess once I manifest my mortality. If I were to exercise my mortality, then (i) I would no longer exist (I take the ceasing of existence to be constitutive of death), and (ii) I would possess F. Given the first-order being constraint, my possessing F entails that I exist. But this means that I both exist (due to possessing F) and do not exist (due to my having died). Since this is impossible, it follows that there is no manifestation condition that I could realize to demonstrate my mortality.

This issue gets its bite when we consider Vetter's argument for

ACTUALITY Potentiality is implied by actuality (2015: 162, 182)

which plays a key role in Vetter's potentiality-based modal semantics. And a central premise in the argument for ACTUALITY is

(2) Whenever an object x exercises a potentiality to Φ, then x must (simultaneously) possess the potential to Φ (2015: 182)

However, the above argument about my mortality suggests that (2) is false. Per (2), if I exercise my potential to die, then I must simultaneously possess the potential to die. But if I possess this potential – which is a property – then, by the first-order being constraint, I exist. So, I must both exist and not exist in order to die. Who knew immortality was so easy!

Of course, there is nothing special about me/my mortality: we can run a similar objection using a variety of objects and potentialities. All that is required is that the relevant potential is both intrinsic and such that manifesting it entails the possessing object's non-existence.

There are several ways that Vetter might respond to this problem. One option is to appeal to extrinsic potentialities.[22] For example, we might say that 'x is mortal' is true because some distinct object *y* has the extrinsic potential for *x* to die. However, this seems like a poor response. First, my mortality seems like one of my intrinsic potentialities, and hence it is part of the ground for some other object's having this extrinsic potential, rather than something that

[22] Thanks to an anonymous referee for pushing me to say more about this response.

it grounded in it. But for me to have this potentiality, I need to be able to manifest it – otherwise, it is not clear in what sense I can be said to have the potential. Second, it is possible to run the same problem using a lonely object – i.e., an object that exists in a world without any other (concrete) objects. Suppose that I am lonely, in the sense that I am the only object that exists in a given world. I will still retain my mortality, and should be able to manifest it. But there is no other object around to possess the relevant extrinsic potentiality. So the move to extrinsic potentialities does not seem to help.

Alternatively, one might suggest that it is possible for me to manifest my mortality, as my doing so does not require my presence. Specifically, as the manifestation of a potentiality is a property, if we held that the manifested property can continue to exist even if the bearer of the property (i.e., me) ceases to exist, then there would be no problem in my manifesting my mortality: the manifested property, sans me, could hang around and do the requisite job.[23]

Obviously, this response depends upon adopting a metaphysics of properties that allows for this kind of "ontologically independent" properties. And there are a number of metaphysical pictures available that do so; for example, one that seems highly suitable for thinking of potentialities in this manner is Tugby's 'Platonic dispositionalism' (2013).

However, this seems like a bad response to the problem. For one, it is not clear how well this conception of properties fits with Vetter's general background object-property ontology (see Vetter 2015: 12-3). More importantly though, on this view, (2) still ends up false. And since (2) is a key part of the argument for ACTUALITY (which is, in turn, a central part of Vetter's modal semantics), the problem remains.

Finally, one could try and block the problem by embracing something like Williamson's *permanentism*, according to which "always everything is always something" (2013: 4). That is, entities never in fact cease to *exist*; instead, when they die/are destroyed, they merely cease to be concrete.[24] Adopting permanentism, the potentialist could say that, when I manifest my potential to die, I simply become non-concrete. And though I am no longer concrete, I still exist, which avoids the contradiction.

This certainly gets around the problem. However, it does so by saddling the potentialist with a pretty heavy – and, to many, counter-intuitive – philosophical commitment. Hence it avoids one cost at the expense of introducing another.

[23] Thanks to an anonymous referee for suggesting this and the next response.

[24] Importantly, being non-concrete is not the same as being abstract. Non-concrete things can possibly be(come) concrete, while abstracta are always and necessarily abstract.

Notably, something like this problem will emerge for any view that, like Vetter's, starts with 'localized' modality (Vetter 2015: 2, though see also Vetter 2010; 2018). Localized views think of possibilities as being 'rooted in objects' – i.e., possibilities are grounded in the objects that are involved in them. Given a localized view, if you accept that (i) I have the potential to die, and (ii) this potentiality entails the truth of, 'I can die', then one is quickly forced into also accepting (iii) for it to be true that I can die, I must be able to bear the property of being dead. For the possibility claim should be grounded in *my* possessing certain properties – that's the appealing part of the localized view after all! But this walks directly into the problem: if I die, I ain't around to bear the property of being dead.

In contrast, non-localized views do not take possibilities to be grounded in the objects they are about. Instead, they are grounded in more "global" matters – for example, in how the various Lewisian worlds are. And, assuming a non-localized view, it is clear that one can accept (i) and (ii) without committing to (iii). For example, one could hold that what makes it true that 'Possibly, I die' is that there is some possible world that has me as a part at an earlier time and does not have me as a part at a later time (because, at some point, I ceased to exist at that world). Importantly, on the non-localized view, something needs to be around to make the possibility (and hence also the potentiality) claim true, but that something doesn't have to be *me*.

The upshot is that this issue isn't a problem for anyone who accepts the truth of 'I can die'. Rather, it's a problem for anyone who thinks that this is a localized possibility that needs to be explained by my manifesting certain properties. Those of us who think it is non-localized can happily accept it without running into trouble.[25]

5. *Conclusion*

The above has raised three issues that emerged from examining the details of Vetter's potentiality-first account of modality. As previously mentioned, these issues are likely not fatal. However, they do highlight potential pitfalls that need to be addressed, as well as places where more clarification would be welcome.

In closing, I would like to stress that, although this paper has raised some issues for the potentiality view, Vetter has offered us a clear, methodically articu-

[25] There is a fourth potential problem facing Vetter's account, that emerges when we consider how her story about degrees of potentiality change over time interacts with POSSIBILITY and the fact that some potentialities will come to be possessed to a maximal degree. However, for space reasons, it is not possible to discuss this matter here. For more of this point, see Wildman (MS).

lated, detailed, comprehensive, and utterly novel account of modality. Though I disagree with several points, her (2015) is an extremely valuable contribution to the broader debate, and a model for how metaphysics should be done.[26]

Nathan Wildman
n.w.wildman@uvt.nl
Tilburg University

References

Anjum, R.L., and Mumford, S., 2018, *What Tends to Be*, Routledge, London.

Bird, A., 2007, *Nature's Metaphysics*, Oxford University Press, Oxford.

Borghini, A. & Williams, N.E., 2008, "A dispositional theory of possibility", in *Dialectica*, 62 (1): 21-41.

Contessa, G., 2016, "Potentiality: From Dispositions to Modality, by Barbara Vetter", in *Mind*, 125 (500): 1236-1244.

Correia, F., 2006, "Generic essence, objectual essence, and modality", in *Noûs*, 40 (4): 753-767.

Correia, F., 2012, "On the reduction of necessity to essence", in *Philosophy and Phenomenological Research*, 84 (3): 639-653.

Fine, K., 1994a, "Essence and modality", in *Philosophical Perspectives*, 8: 1-16.

Fine, K., 1994b, "Senses of essence", in W. Sinnott-Armstrong, D. Raffman, and N. Asher (eds.), *Modality, morality and belief. Essays in honor of Ruth Barcan Marcus*, Cambridge University Press, Cambridge.

Fine, K., 1995, "The logic of essence", in *Journal of Philosophical Logic*, 24 (3): 241-273.

Fine, K., 2000, "Semantics for the logic of essence", in *Journal of Philosophical Logic*, 29 (6): 543-584.

Jacobs, J.D., 2010, "A powers theory of modality: or, how I learned to stop worrying and reject possible worlds", in *Philosophical Studies*, 151 (2): 227-248.

Jenkins, C.S., and Nolan, D., 2012, "Disposition Impossible", in *Noûs*, 46 (4): 732-753.

Kimpton-Nye, S., 2018, "Hardcore Actualism and Possible Non-Existence", in *Thought*, 7 (2): 122-131.

Lange, M., 2009, *Laws and Lawmakers Science, Metaphysics, and the Laws of Nature*, Oxford University Press, Oxford.

Leech, J., 2017, "Potentiality", in *Analysis*, 77 (2): 457-467.

Lewis, D.K., 1986, *On the Plurality of Worlds*, Wiley-Blackwell, Hoboken NJ.

Manley, D., and Wasserman, R., 2008, "On linking dispositions and conditionals", in *Mind*, 117: 59-84.

[26] I would like to thank Amanda Cawston, Umut Baysan, and Stephan Leuenberger for discussing several of the ideas in this paper, and, of course, Barbara Vetter for her exciting and interesting ideas!

McKitrick, J., 2019, "Barbara Vetter: Potentiality: From Dispositions to Modality", in *Erkenntnis*, 84: 1179-1182. DOI: https://doi.org/10.1007/s10670-018-9999-2

Pruss, A., 2002, "The actual and the possible", in R.M. Gale (ed.), *Blackwell Guide to Metaphysics*, Blackwell, Oxford: 317-33.

Schrenk, M., 2015, "Trigger Happy. Ein Kommentar zu Barbara Vetters Potentiality", in *Zeitschrift für Philosophische Forschung*, 69 (3): 396-402.

Tugby, M., 2013, "Platonic dispositionalism", in *Mind*, 122 (486): 451-480.

Vetter, B., 2014, "Dispositions without Conditionals", in *Mind*, 123 (489): 129-156.

Vetter, B., 2015, *Potentiality: From Dispositions to Modality*, Oxford University Press, Oxford.

Vetter, B., 2016, "Counterpossibles for dispositionalists", in *Philosophical Studies*, 173 (10): 2681-2700.

Wildman, N., ms., *Potential contradictions*, unpublished manuscript.

Williamson, T., 2013, *Modal Logic as Metaphysics*, Oxford University Press, Oxford.

Potentiality, modality, and time

Jennifer Wang

Abstract: Barbara Vetter's project in *Potentiality* is to articulate and defend a dispositionalist theory of modality based on potentialities. My focus is on the metaphysics of her positive theory. I consider one of Vetter's main targets, David Lewis's theory of possible worlds, and use it to distinguish what I call "*de re* first" approaches from "*de dicto* first" approaches. This way of framing the disagreement helps shed light on what their respective accounts can intuitively accomplish. In particular, I introduce objections to Vetter's requirement that the grounds of *de dicto* modal truths must be routed through time. I also suggest an alternative *de dicto* first approach that Vetter does not consider, one which does not come saddled with Lewis's ontology or with Vetter's issues with *de dicto* modal truths. Rather, on incompatibilism, modality is grounded on second-order relations between (non-potentialist) properties, e.g. incompatibility or entailment. Defenders of *de dicto* first approaches, including incompatibilism, can better account for such *de dicto* modal truths, thus undermining some of the intuitive appeal of Vetter's theory.

Keywords: potentiality; modality; dispositions; possible worlds; properties; incompatibility.

> "Possible chicken means actual egg – plus actual sitting hen, or incubator, or what not".
> William James, *Pragmatism*

1. *Introduction to the project*

Barbara Vetter's project in *Potentiality* (2015) is to articulate and defend a dispositionalist theory of modality based on potentialities. Potentialities are properties of individual objects often expressed by terms in English such as "fragility" (the potentiality to break) and "irascibility" (the potentiality to get angry). Some potentialities are classified as dispositions; but whereas dispositional ascriptions are context-sensitive and vague – features of language, not the world – potentialities are properties belonging to the underlying metaphys-

philinq VIII, 1-2020, pp. 185-198
ISSN (print) 2281-8618-ETS DOI: 10.4454/philinq.v8i1.281

ics. Vetter argues forcefully that potentiality is the more basic theoretical notion, one that encompasses not only cases of dispositions, but also abilities and powers. For Vetter, potentialities are (i) individuated by their manifestations alone, and are (ii) linked to possibility. So whereas defenders of what she calls the "standard conception" of dispositions typically base their theories of modality on counterfactual conditionals, Vetter bases hers on actual potentialities of concrete individuals.

Vetter's defense of a potentiality-based modal theory is truly impressive. In the course of developing this theory, she provides a novel logic and linguistic semantics, something that few other defenders of modal theories even attempt. I will have relatively little to say about these tasks. My focus will be on the metaphysics of her positive theory.

I will say more about Vetter's theory of modality after further introducing her notion of potentiality. I will consider one of Vetter's main targets, Lewis's (1986) theory of possible worlds, and use it to distinguish what I call "*de re* first" approaches from "*de dicto* first" approaches. This way of framing the disagreement helps shed light on what their respective accounts can intuitively accomplish. In particular, I will introduce objections to Vetter's theory's requirement that the grounds of *de dicto* modal truths must be routed through time.[1] I also suggest an alternative *de dicto* first approach that Vetter does not consider, one which does not come saddled with Lewis's ontology or with Vetter's issues with *de dicto* modal truths. Rather, on incompatibilism, modality is grounded on second-order relations between (non-potentialist) properties, e.g. incompatibility or entailment. Defenders of *de dicto* first approaches, including incompatibilism, can better account for such *de dicto* modal truths, thus undermining some of the intuitive appeal of Vetter's theory.

2. *Potentiality and modality*

This section introduces the various pieces required to understand the role of potentialities in Vetter's theory of modality, which is based on the following principle:

POSSIBILITY: It is possible that $p =_{df}$ Something has an iterated potentiality for it to be the case that *p*.[2]

[1] Note that following Vetter, I will make free use of grounding talk, in addition to the more traditional notion of reduction; see her 1.6.

[2] For similar ideas, see especially Borghini and Williams (2008), Contessa (2010), and Pruss (2002).

Potentialities come in degrees. Consider the disposition term "fragility", which is context-sensitive. In normal circumstances, we would consider a champagne glass to be fragile, and to be more fragile than a tumbler. Both in turn would be more fragile than a diamond, which is ordinarily not considered fragile at all. And between the tumbler and the diamond are cases where it is not clear whether we should say the object in question is fragile or not – that is, there will be vagueness in the cut-off point between fragile and non-fragile. But for Vetter, the source of context-sensitivity and vagueness is tied to language rather than the world. Metaphysically speaking, the champagne glass, tumbler, and diamond all share the potentiality to break, but they each possess this potentiality to a differing degree.

Potentialities may be possessed to the *maximal* degree. In such cases, if something has the potentiality to F, then it must F (and thus lacks the potentiality not to F). A massive object always attracts other massive objects – it could not do otherwise. Conversely, potentialities may be possessed to a very low degree. Anything that can break thereby has the potentiality to break, unlikely though this may be.

There are joint, extrinsic, and iterated potentialities. So far, we have only considered cases where an individual object has the potentiality to F. Vetter holds that we should also accept cases where some objects jointly have the potentiality to F. An example she uses throughout is of a key and the door that it unlocks. The key alone has the potentiality to open locks of a certain shape; the door alone has the potentiality to be opened by keys of a certain shape. Together, they have the potentiality to stand in the relation of opening – that is, they have the potentiality for the key to open the door. For another example, the people in a crowd have the potentiality to stampede, though no single one of them does alone. The manifestation of joint potentialities falls into one of three categories: a relation between individuals, a plural property, or an individual property.

Intrinsic potentialities are properties that are intrinsic to their bearers; likewise, intrinsic joint potentialities concern only the plurality of objects that possess them. In contrast, extrinsic potentialities concern objects extrinsic to their bearers. For instance, the key has an extrinsic potentiality to open a particular door. (However, interestingly enough, the key also has an intrinsic potentiality to open doors whose locks have a particular shape.) Vetter is liberal about the existence of extrinsic potentialities. She argues that anytime some objects possess a joint potentiality, each individual object possesses a corresponding extrinsic potentiality.[3]

[3] She also holds that any time an object possesses an extrinsic potentiality, the object together with other objects have a corresponding joint potentiality.

There are iterated potentialities: potentialities to acquire potentialities (which may themselves be potentialities to acquire potentialities, etc.). The addition of iterated potentialities allows the theory to extend its "reach", as Vetter puts it. Consider her case of the possibility that she plays the violin. Vetter currently does not have the ability to play the violin, but has the ability to learn how to play the violin; thus, she has an iterated ability to play the violin. This is a twice-iterated potentiality, in contrast with the once-iterated potentialities we have thus far been considering. Furthermore, some violin teacher has an ability to enter into a joint potentiality with Vetter for Vetter to learn to play the violin. This is a three-times iterated potentiality.

We now have enough theory in place to understand Vetter's modal principle POSSIBILITY. The basic picture is this. If *p* is possible, then there will be some objects which jointly have an (*n*-)iterated intrinsic potentiality for *p*.[4] The *p* in question may express a relation between the objects or an individual property of one of the objects. But as noted, any joint potentiality can be expressed as an extrinsic potentiality of an individual object, and thus POSSIBILITY should be extensionally adequate.

We have arrived at a theory of modality that locates the source of possibility claims in the properties of actually existing, concrete individuals, which will be appealing to many. Furthermore, what's possible has to do with the actual history of the universe and the different ways it could have unfolded. For something to have the potentiality to be such that *p*, that manifestation must lie in that thing's present or possible future. Vetter writes (2015: 186): "it is true of me now that I was once a child, but it would be odd to say that I now have a potentiality to have been a child". Thus, she holds that potentiality is "forward-looking" in time. I will discuss unintuitive consequences of the interaction of potentiality and time in section 4. But first, I'd like to consider some consequences of Vetter's theory of modality as contrasted with one of its main rivals, possible worlds theory.

3. De re *first vs.* de dicto *first modality*

Vetter acknowledges opponents on two sides when it comes to modal theorizing: the possible worlds theorist and the traditional dispositionalist. On the one hand, she argues against dispositionalist theories of modality that begin with a purported link between dispositions and counterfactual conditionals.

[4] Notice that *p* is a proposition rather than a property. Vetter holds that this is an innocent construction for expressing manifestations, e.g. the potentiality for the door to open versus the potentiality that the door opens. See Vetter (2015: 104).

On the other hand, she argues against possible worlds theory, most notably modal realism as defended by Lewis (1986). In the remainder of this paper, I will argue that Vetter's view faces deeper problems than she recognizes. Furthermore, such problems are avoided not only by Lewis's view, but by another view that bases modality in actually existing properties. That there is conceptual space for this view is made clear by classifying views as "*de re* first" or "*de dicto* first"; the former ground all modal facts in *de re* modality, the latter in *de dicto* modality. This classification will be explained and explored in this section.

As Vetter notes, potentiality and dispositionality are among a cluster of related modal notions which include essence, counterfactuality, causation, and possibility (and its dual, necessity). She classifies potentiality as a "localized modality", as potentialities are properties of individuals, and should be construed formally as a predicate operator. In contrast, she classifies possibility as a "non-localized modality", one which need not concern particular individuals. Here, she is thinking of possibility construed formally as a one-place sentential operator.[5] Vetter differentiates her localized/non-localized distinction from the more familiar *de re/de dicto* distinction; while the latter applies to sentences, the former is "straightforwardly metaphysical" (3, footnote 3).[6] Furthermore, there are cases of *de re* possibility (or necessity) claims that are not potentiality (or essence, the dual of potentiality) claims.[7]

Nonetheless, in proceeding, I will frame the debate between Vetter and relevant opponents in terms of whether they are "*de re* first" or "*de dicto* first" views. While it is true that the *de re/de dicto* distinction is typically applied to sentences (or propositions), it is now generally understood by metaphysicians to capture a difference in the world as well: *de re* modality concerns the modal properties of particular individuals, whereas *de dicto* modality concerns purely general possibilities or necessities.[8] Potentialities are *de re* modal properties, even if not all *de re* modal claims correspond to potentiality claims.

[5] See her sections 1.1-2 for discussion. I will not take a stand on the proper formalism for expressing various modal notions (though Vetter does a thorough job of arguing that potentiality should be formalized as a predicate operator in her chapter 5). What matters is that Vetter rejects the well-known possible worlds framework for understanding the metaphysical basis of modal claims. Note that she is fine with using possible worlds talk for instrumental purposes; see her chapter 3.

[6] A *de re* modal claim, when formalized in terms of sentential operators, is one where a modal operator either has a free variable or a name in its scope.

[7] Vetter offers two examples: (i) while it is *de re* possible that she not exist, she does not thereby have a potentiality to not exist (such a possibility being instead grounded in potentialities of her parents) (194); and (ii) while it is necessary that Socrates belongs to his singleton set, it is not essential to him (3fn3).

[8] See Nelson (2019) for an explanation of the more traditional syntactic or semantic distinctions, as well as the metaphysical one.

Lewis's modal realism is a *de dicto* first view. His concrete possible worlds ground possibility claims via this biconditional: It's possible that *p* if and only if there exists a possible world in which *p*. But this is only the start of a systematic reduction of the notion of possibility. The biconditional straightforwardly holds for *de dicto* modal claims – but Lewis also needs a way to make sense of *de re* modal claims, since individuals are worldbound on his view. His solution is counterpart theory. According to counterpart theory, a *de re* modal claim like "I could have had a sister" are true in virtue of my having a counterpart in some world that has a sister. This counterpart is relevantly similar to me, where what counts as relevantly similar is supplied by context. *De re* modality is thus reduced to *de dicto* modality on Lewis's view.

Vetter, in turn, wants to base possibility in the dispositional properties of actually existing objects. The right-hand side of POSSIBILITY is a *de re* modal claim; it is hence easy to see how Vetter accounts for *de re* possibility claims. But there aren't many places in her book where Vetter explicitly, directly addresses the question of how to ground *de dicto* possibility claims. In her section 6.2, she briefly considers three cases of *de dicto* possibilities (202):

(1) It is possible that there is a woman president of the US.
(2) It is possible that there be a human space station on Mars.
(3) It is possible that humans should have three legs instead of two.

POSSIBILITY requires a witness for any possibility claim: for *p* to be possible, some individual in the history of the world (past to present) must have an iterated potentiality for it to be case that *p*. Any woman who at any point had the potentiality to be president of the US is thereby is a witness for (1). A witness for (2) is any engineer who at any point had the potentiality to be among a team that builds a human space station on Mars; this potentiality will be extrinsic and based on a joint potentiality of the entire team. The tricky case is (3). To find a witness, says Vetter, we have to look at our pre-human ancestors, who had an iterated potentiality to have offspring that is human. If such pre-human ancestors also had an iterated potentiality to have offspring that is human and three-legged, then we have found witnesses for (3). Thus, Vetter's theory locates the source of this *de dicto* possibility in the *de re* by "rewinding" time and looking at the potentialities of our ancestors. Similar considerations hold for the possibility of there being talking donkeys, unicorns, etc.

These cases highlight the intuitive appeal of Vetter's view, especially in comparison with Lewis's view. For her, all possibilities require a basis in the actually-instantiated potentialities of past- or presently-existing individuals. This should be appealing to those who are suspicious of the existence of other possible worlds. But modal realism is not the only possible worlds theory. There are other

de dicto first approaches that accept possible worlds as actually-existing abstract objects. For instance, Vetter discusses Stalnaker's (1976; 2003) view on which possible worlds are properties that the world could have instantiated.[9] According to Stalnaker, *p* is possible just in case if some world-property were instantiated, then *p* would be true. Vetter argues that her own theory is preferable because it locates the source of modality at a lower level (2015: 265): "[T]he world […] has a potentiality to be such that I am sitting. However, the world has that potentiality in virtue of *my* having the potentiality to be sitting, not vice versa".

But someone sympathetic to the idea of world properties is also free to locate the source of modality at a lower level. The fact that some world property *w* could be instantiated need not be a brute fact; it may be grounded in the co-instantiability of the more local properties involved. Importantly, this need not require potentialities at all. The incompatibilist posits primitive incompatibilities between certain properties, so that what's possible or not comes down to whether or not the properties involved are compatible. This does not require positing primitive incompatibilities between any two properties that are incompatible – some properties may be derivatively incompatible in virtue of their relations to other properties. For instance, being a square circle is an uninstantiable property in virtue of the incompatibility of the properties of being square and being circular. Incompatibilism requires a story of how derivatively incompatible properties are related to primitively incompatible properties, but this story may be filled out in different ways.[10]

Incompatibility is one of a cluster of notions relating properties and relations. Two properties are compatible just in case they are not incompatible. And entailment between properties may be defined as so: P entails Q just in case P is incompatible with not-Q. As such, the choice of incompatibility as the modal primitive is arbitrary. The incompatibilist could equally well take compatibility or entailment as her primitive modal notion and define the other notions accordingly.[11] I will appeal to any of these notions when discussing the basis of modal truths according to the incompatibilist.

[9] Since Stalnaker's theory appeals only to properties that *could* be instantiated rather than those that cannot, it is not a reductive theory of modality – but neither is Vetter's theory.

[10] Defenders of versions of incompatibilism include Jubien (2007; 2009), Lycan (1994), and Wang (2013). Bigelow and Pargetter (1990) and Forrest (1986a, 1986b) defend views closely related to incompatibilism, though they prefer to avoid primitive modality.

[11] One may wonder whether these are really primitive modal notions. After all, property P entails property Q iff necessarily, if object x instantiates P, then x instantiates Q. Doesn't this show that the notion of entailment between properties is analyzable in terms of possibility and necessity, so that incompatibilism collapses into a modalist view? This objection gets the direction of dependence the wrong way around: the incompatibilist insists that the right-hand side of the biconditional is explained by the left-hand side.

Notice that I am counting incompatibilism as a *de dicto* first approach to modality, despite the fact that it crucially appeals to properties. This is because the source of modality is not the instantiation of modal properties by actual, existing objects, but rather, the modal relations between the properties themselves. And as long as the incompatibilist posits primitive incompatibilities between only general properties, it is primarily about *de dicto* modality. Of course, the incompatibilist also requires an account of *de re* modality. She is free to posit primitive incompatibilities between individual-involving properties, resulting in neither a *de re* nor *de dicto* first approach to modality, or to appeal to something like counterpart theory.[12]

4. *Potentiality and time*

I argue in the remainder of this section that a theory like incompatibilism has an advantage over Vetter's theory: unlike Vetter's theory, it intuitively accounts for *de dicto* modal claims that (i) should not or (ii) cannot depend upon past or present individuals. I do not consider the problem cases presented below to be counterexamples to Vetter's theory; after all, she is willing to bite the bullet about various consequences of her view. Rather, I take them to undermine a crucial selling point of her theory: its intuitive attractiveness. Vetter emphasizes her theory's intuitive appeal at various points in her book – and I agree that it has this appeal when it comes to some *de re* modal truths.[13] But if her theory loses its intuitive appeal when it comes to certain *de dicto* modal truths, then she cannot claim a clear advantage over her *de dicto* first rivals.

First, there are cases of *de dicto* modal claims that should not depend upon past or present individuals. For Vetter, the basis of any modal claim is fundamentally diachronic, specifically, is past- or present-involving.[14] But there are cases that do not seem to involve time at all. Consider the sorts of cases that motivate incompatibilism in the first place. Jubien's (2009: 92) examples include: (i) the property of being square entails the property of having linear sides; (ii) the property of being yellow entails the property of being colored; and (iii) the property of being a spouse entails the property of being married; and (iv) the property of being a horse entails being an animal. These are

[12] See Wang (2015a).

[13] See for instance (11), where she talks about the naturalness of her ontology, and (14 footnote 14), where she mentions the intuitive appeal of her theory.

[14] It could be future-involving as well. However, given Vetter's views, for a future object to be potentially such that *p*, there must be a past or present object that has an iterated potentiality to be such that *p*. So for ease of discussion, I will stick with the more intuitive characterization of her view as past- or present-involving.

the basis for the following modal truths: (i) necessarily, all squares have linear sides; (ii) necessarily, all yellow things are colored; (iii) necessarily, all spouses are married; and (iv) necessarily, all horses are animals. Thus, the source of the necessity of all squares having linear sides is located in an entailment relation between the property of being a square and the property of having linear sides, and so on for the rest.

There may, of course, be disagreement over whether the incompatibilist is correct about the basis of these modal truths. One may consider these examples as mere analyticities (though I do not). To sidestep this possible distraction, let's focus on an example inspired by Wang (2013):

(4) Necessarily, no negatively charged objects are positively charged.

For the incompatibilist, the source of this truth is an incompatibility relation between the property of being negatively charged and the property of being positively charged. But Vetter must say that the source is past- or present-involving. For Vetter, necessity is defined as the dual of possibility (203):

NECESSITY: It is necessary that p $=_{df}$ It is not possible that not-p.

This works out to the view that it is necessary that p iff nothing has an iterated potentiality to be such that not-p. In this case, nothing in the history of the universe has an iterated potentiality to be such that something is both negatively and positively charged. But why should that be the case, since this does not involve a logical impossibility?[15]

It may be helpful to approach the question from the other direction. Recall from above that potentialities may be possessed to the maximal degree. Thus, something that is maximally fragile must shatter. This only allows us to attribute necessary properties to an individual, that is, make the following *de re* modal claim: *x* must be fragile. Vetter may try to find a way to get from *de re* attributions of necessity to *de dicto* necessity claims. Perhaps she could say in the case of (4) that anything that is negatively charged must possess the potentiality to be non-positively charged to the maximal degree. This will get her: All negatively charged objects are necessarily not positively charged. But this is still only a *de re* modal predication rather than a *de dicto* claim.

For another example along these lines, consider what the incompatibilist would say about *de dicto* possibilities such as: "There could exist a red square". For the incompatibilist, this is true in virtue of the compatibility of the property of being red and the property of being square. There is no need to say of

[15] One may replace this example with their favorite example of incompatible but non-logically-contradictory properties, such as being red and being blue, as needed.

any actually existing thing that it has the potentiality to be such that there is a red square. For Vetter, any presently-existing red square serves as the witness for this claim. But before there were any red squares, there were still objects that had an iterated potentiality to be such that there is a red square. Presumably, the "best" witness would be some object that along with other objects had the joint potentiality to produce a red square. All this may sound fine. However, Wang (2015b) points to another kind of case: consider the apparent possibility that a glass appears *ex nihilo*. There is no logical impossibility, and hence, according to many, no metaphysical impossibility involved in this supposition. Vetter must either deny that such a case is possible, or try to find a plausible witness. I don't think that she would go for the latter; what would be witness be? One might be tempted to say that the witness would be the world. However, in her section 7.4, Vetter argues that even if the world is a bearer of potentialities, it bears potentialities in virtue of the potentialities of "smaller" individuals, e.g. a subregion of the world. But the potentialities in a subregion can only guarantee that there is no glass appearing *ex nihilo* in that region, rather than globally.[16]

I turn now to cases of *de dicto* modal claims that cannot depend upon past or present individuals. Consider the apparent possibility (also discussed in Wang 2015b) that a glass always exists in a universe with no beginning, or more simply:

(5) It's possible for there to be an object that always exists in a universe with no beginning.

For Vetter, such an object is not possible; in fact, such a universe is not possible. For non-actual possibilities must have their basis in past or present actual objects, and no actual object has a potentiality for an object to have always existed in a universe with no beginning. (Note: Vetter does assume that the universe actually has a beginning, but even if she didn't, there would still not be any object in our past that could be the basis of the truth of (4).)

Vetter discusses similar cases as potential counterexamples, but writes (290): "The intuition that there could always have been different objects is, I believe, not at the centre of our modal intuitions, and like many philosophical intuitions it may well be theory-driven. The very same temporal asymmetry was shown to explain a modal principle that is, I believe, more central and accepted by many philosophers: the necessity of origin. Indeed, this further consequence of temporal asymmetry can be seen as another application of the necessity of origin. It is the origin of the universe itself, in precisely the objects

[16] Thanks to an anonymous referee for pressing me on this point.

which originally constituted it, that is necessary on the present view". I will not review Vetter's reasons in favor of the necessity of origins (204-6), as I think that the resulting view is still problematic (as explained below). Eternal glasses are outlandish, but universes with different origins or no origins are not.

In her section 7.9, Vetter recognizes that the necessity of the origin of the universe is a controversial thesis, but aims to assuage some worries by arguing that this does not imply the necessity of actuality. The worry that Vetter addresses goes something like this: if the beginning of the universe is necessary – that is, the first total state of the universe is necessary – and if the laws of nature are deterministic, then all following states will also be necessary. But Vetter points out that her view is not committed to this. It may be that there is only one possible future at a time when taking into account the state of the total universe. However, POSSIBILITY allows for individuals in a subregion of the universe to ground different possible futures in virtue of their joint iterated potentialities – it's just that those potentialities will be frustrated by the potentialities of individuals outside that subregion.

This is effectively the reasoning behind Vetter's denial that metaphysical modality just is nomic modality. In her section 7.8, Vetter favors a "best systems" account of the laws of nature (though she also thinks other theories of laws are compatible with her metaphysics).[17] The laws of nature will be the best systematization of the distribution of fundamental properties, including potentialities. Even if one thinks that all fundamental properties are potentialities, different systematizations are possible depending on which potentialities are instantiated by actual, concrete objects. As I think a thorough treatment of this interesting idea deserves more space than I can give it here, I will simply accept Vetter's reasoning. My focus is on the unintuitive consequence that remains: Vetter cannot accept the possibility that the universe had a different origin or no origin at all. And this will limit what she regards as nomically possible.

In summary, (4) and (5) reveal that there is something lacking in Vetter's *de re* first approach compared to a *de dicto* first approach, because of the inappropriate involvement of time on Vetter's theory. The worry for Vetter with the first kind of case is that dispositionality is a diachronic notion, and is thus ill-suited to account for synchronic modal truths. The second kind of case reveals that there are possibilities that do seem to concern the past or present, but do not seem to concern *our* past or present. In contrast, *de dicto*

[17] For a best systems account that takes the fundamental properties to be potentiality-like (specifically, powers), see Demarest (2017). Demarest (2015) contains an overview of the relevant theories of laws of nature.

first approaches say that such cases are possible just in case there are possible worlds in which they are actual (Lewis), or just in case the properties involved are compatible with each other.

5. *Concluding remarks*

In this paper, I have aimed to provide a concise summary of Vetter's theory of modality along with some critical remarks. I have argued that although Vetter's theory has intuitive plausibility in many cases of *de re* modal claims, it is counterintuitive in at least some cases of *de dicto* possibilities. On this front, competitors who have a *de dicto* first approach to modality fare better. This undermines the intuitive plausibility that Vetter claims for her view.

I end with a suggestion. Vetter may wish to consider adopting a hybrid view on which potentialities account for some modal truths, while primitive incompatibilities account for others. This would allow her to locate the source of modality in properties, which would still be an attractive alternative to Lewis's modal realism for many. But she could not say that the properties of actual, concrete objects account for all modal truths (since in some cases, the properties themselves and the relations between them would serve as the basis). Still, this may be a better path forward for Vetter than to bite the bullet.

Jennifer Wang
jwa265@sfu.ca
Simon Fraser University

References

Bigelow, J., Pargetter, R., 1990, *Science and Necessity*, Cambridge University Press, Cambridge.

Borghini, A., Williams, N.E., 2008, "A Dispositional Theory of Modality", in *Dialectica*, 62, 1: 21-41.

Bricker, P., 1996, "Isolation and Unification: The Realist Analysis of Possible Worlds", in *Philosophical Studies*, 84, 2-3: 225-238.

Bricker, P., 2001, "Island Universes and the Analysis of Modality", in G. Preyer and F. Siebelt (eds.), *Reality and Humean Supervenience: Essays on the Philosophy of David Lewis*, Rowman and Littlefield, Lanham MD: 27-56.

Contessa, G., 2010, "Modal Truthmakers and Two Varieties of Actualism", in *Synthese*, 174, 3: 341-353.

Demarest, H., 2015, "Fundamental Properties and the Laws of Nature", in *Philosophy Compass*, 10, 5: 334-344.

Demarest, H., 2017, "Powerful Properties, Powerless Laws", in J. Jacobs (ed.), *Causal Powers*, Oxford University Press, Oxford: 38-53.

Divers, J., 2002, *Possible Worlds*, Routledge, London.

Forrest, P., 1986a, "Neither Magic nor Mereology: A Reply to Lewis", in *Australasian Journal of Philosophy*, 64, 1: 89-91.

Forrest, P., 1986b, "Ways Worlds Could Be", in *Australasian Journal of Philosophy*, 64, 1: 15-24.

Jacobs, J., 2010, "A Powers Theory of Modality: or, How I Learned to Stop Worrying and Reject Possible Worlds", in *Philosophical Studies*, 151, 2: 227-248.

Jubien, M., 1996, "Actualism and Iterated Modalities", in *Philosophical Studies*, 84, 2-3: 109-125.

Jubien, M., 2007, "Analyzing Modality", in D. Zimmerman (ed.), *Oxford Studies in Metaphysics volume 3*, Oxford University Press, Oxford: 99-139.

Jubien, M., 2009, *Possibility*, Oxford University Press, Oxford.

Lewis, D., 1986, *On the Plurality of Worlds*, Wiley-Blackwell, Oxford.

Lycan, W., 1994, *Modality and Meaning*, Kluwer Academic Publishers, Dordrecht.

McDaniel, K., 2004, "Modal Realism with Overlap", in *Australasian Journal of Philosophy*, 82, 1: 137-152.

Nelson, M.. 2019, "The De Re/De Dicto Distinction", in E.N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2019 Edition), <https://plato.stanford.edu/entries/prop-attitude-reports/dere.html>.

Pruss, A., 2002, "The Actual and the Possible", in R.M. Gale (ed.), *The Blackwell Guide to Metaphysics*, Blackwell, Oxford: 317-333.

Stalnaker, R., 1976, "Possible Worlds", in *Noûs*, 10, 1: 65-75.

Stalnaker, R., 2011, *Mere Possibilities*, Princeton University Press, Princeton.

Vetter, B., 2014, "Dispositions Without Conditionals", in *Mind*, 123, 489: 129-156.

Vetter, B., 2015, *Potentiality*, Oxford University Press, Oxford.

Wang, J., 2013, "From Combinatorialism to Primitivism" *Australasian Journal of Philosophy*, 91, 3: 535-554.

Wang, J., 2015a, "Actualist Counterpart Theory", in *Journal of Philosophy*, 112, 8: 417-441.

Wang, J., 2015b, "The Modal Limits of Dispositionalism", in *Noûs*, 49, 3: 454-469.

Replies

Barbara Vetter

Abstract: This paper responds to the contributions by Alexander Bird, Nathan Wildman, David Yates, Jennifer McKitrick, Giacomo Giannini & Matthew Tugby, and Jennifer Wang. I react to their comments on my 2015 book Potentiality: From Dispositions to Modality, and in doing so expands on some of the arguments and ideas of the book.

Keywords: dispositions; powers; potentiality; modality.

I am very grateful to Lorenzo Azzano and Andrea Borghini, and to the seven contributors for this special issue, from whose papers I have learned a great deal. I am deeply honoured (and quite delighted) by the time and effort they invested in thinking about potentiality in general, and about *Potentiality* in particular. In what follows, I will not be able to address every point in as much detail as it deserves, and in some cases I will only indicate the direction in which I think further discussion should go. These are not rhetorical devices, but rather expressions of my hope that these discussions will continue.

1. *Dispositions and conditionals: response to Bird*

Potentiality, like many metaphysical topics, can be approached from two directions. We can come to it from 'the manifest image', our everyday understanding of our own abilities and the dispositions of the objects in our environment. Or else, we can approach it from 'the scientific image', taking our best scientific theories, perhaps interpreting them and drawing whatever conclusions can be drawn from them. There is no guarantee that the two images will converge; often, it is argued that they are in opposition to each other. One tenet of my book is that in the case of potentiality, the two images happily converge on a common picture – and in fact that they do so twice over. First, both images make it natural to take potentiality as a primitive, not to be reduced away as it would be in a Humean ontology. In this, I believe, Alexander Bird and I

ISSN (print) 2281-8618-ETS DOI: 10.4454/philinq.v8i1.282

are in agreement. Second, both images suggest a conception of potentiality or – to use the more common term from which I start – of dispositions that is very different from the conception which philosophers have standardly accepted. It is here that Alexander Bird disagrees with me.

In chapters 23 of *Potentiality*, I rejected the orthodox conception which links dispositions, either reductively or non-reductively, to counterfactual conditionals, and developed instead a conception of dispositions that links them to graded possibility. To characterize ordinary dispositions such as fragility, instead of

(F-S) *x* is fragile iff were *x* subjected to a stress, *x* would break,

I proposed

(F-V') *x* is fragile iff *x* could break easily.

My initial argument for (F-V') included both semantic and metaphysical considerations, semantics being, I take it, one of our best guides to the 'manifest image'. I then turned to the scientific im age to discuss "nomological dispositions" (ch. 2.6), i.e., dispositions that encode a genuine law of nature, and argued that for non-semantic reasons they, too, are not best captured by a conditional conception. (The argument is a generalization of my criticism of Bird 2007, as formulated in Vetter 2012.) Rather, I argued, nomological dispositions are best integrated into the picture that arose from the manifest image: potentiality, which is characterized only by a manifestation, comes in degrees; ordinary dispositions such as fragility are situated at the lower end of the degree spectrum, while nomological dispositions are to be found at its higher, maximal end.

I did not then and I do now take myself to be "drawing conclusions about fundamental aspects of modality from evidence regarding our use of everyday expression" (Bird: 95). Semantic considerations about everyday expressions illuminate the manifest image, and the concepts we use even in understanding the scientific image. And being clear on our concepts is useful even when we go on to apply them to understand the scientific image. But it is considerations about nomological dispositions themselves, not about semantics, that justify my including them in the picture which is suggested by the semantics.

So much for methology. Bird, however, has objections both to my claims about semantics (the manifest image) and about nomological dispositions (the scientific image). I believe that I can dispel the first objection, but the second will require more extensive discussion than I can provide here (and now).

Let me take up the semantic objection first, then. With regard to (F-V'), Bird asks: "How is the 'easily' qualification supposed to be understood?", and

answers: "It is to be understood, as others have done, in terms of close possible worlds" (Bird: 88), i.e. as *x* breaking in at least one (or a few) of the possible worlds that share our laws and have "exact matches in matters of particular fact" (Bird: 88 fn 2). He then goes on to provide a counterexample to (F-V') thus understood, where a fragile glass is from its creation so shielded that it does not break in any close world, i.e., its fragility is always masked. His diagnosis is that I have mistaken the force of 'easily', which he thinks qualifies the implicit stimulus and not the nature of the modality involved.[1]

My response is different from the one that Bird envisages (89-90), however. For I do *not* appeal to closeness as understood here. Rather, I argue that the easy possibility involved in dispositions is best captured by a *proportional* model, if it is captured in possibleworlds terms at all: whether or not an object *x* is fragile is a matter of *x*'s breaking in a *sufficient proportion* or relevant worlds (Vetter 2015: 72ff.). Those relevant worlds, in turn, are not supposed to be the close worlds of Lewisian semantics. Rather, they should "provide maximal variation in the external circumstances. The proportion of cases in which a vase breaks … should not depend on factors that are external to the vase" (Vetter 2015: 77; see also Vetter 2014).[2] Bird's shielded glass, on this understanding, will break in as many of the relevant worlds (or cases) as one that is not so shielded, because its being shielded will not be held fixed across the relevant worlds. (For more on the proportional understanding of dispositions, see sections 2 and 3, and the contributions by Wildman and Yates to which they respond.) This, I believe, rebuts Bird's first worry.

Bird's second worry, however, is not so easily dispelled. So let us turn to that worry: the "problem of non-conditional possibilities".

The problem arises from three assumptions (together with some observations on modal logic):

1. On a dispositionalist view all modality has to depend (or, Bird says, supervene) on dispositions at the fundamental level.
2. All dispositions that are fundamental are nomological dispositions (i.e. dispositions that encode a law of nature).
3. All nomological dispositions have an internal conditional structure.

[1] Note that Lowe (2011) and Aimar (2018) would both agree with Bird on the reading of 'easily' and still opt for a possibility account of dispositions. I will not take that route, but I think it is a live option.

[2] Note that this applies, as it stands, only to intrinsic dispositions. For extrinsic dispositions, those external factors on which the disposition depends should not be varied; see Vetter 2015, 75f. fn10.

The first assumption arises simply from the idea that everything is grounded in the fundamental level; I will not question it in what follows. The second assumption is natural given that we take fundamental physics to be our best guide to the fundamental level. The third assumption is one which I shared in Vetter 2015, chs. 2.5-2.6 and 3.5; what I argue there is that the conditional nature of nomological dispositions is better characterized by my view of dispositions, which takes the conditional nature to be embedded in the disposition's manifestation, rather than giving the disposition itself the structure of a (counterfactual) conditional. Given the three assumptions, however, the fundamental dispositions are all conditional: they give rise to possibilities (or indeed necessities) for certain conditionals to be true, but they do not yield any non-conditional possibilities.

The result is somewhat ironic since my view was precisely characterized by its focus on possibilities and *not* conditionals; accordingly, the problem starts out in Bird's paper as a problem for the standard conception of dispositions and an advantage for my view. Integrating the laws of nature into the picture, however, appears to force conditionals back onto us and thus makes the problem raise its head even within the alternative conception of dispositions. What this makes clear, however, is that the problem is not a specific problem for my (non-conditional) account of dispositions; it is rather a problem for dispositionalism about modality quite generally.

I do not have the space here to provide a full solution to this interesting problem, nor am I certain what a solution would look like. But an obvious place to look for a solution is in the assumptions that I have made explicit above.

Consider assumption 2: Are all fundamental dispositions nomological dispositions? This does not seem obvious to me. Even if physics supplies nothing but nomological dispositions, and (part of) physics is our best science of the fundamental, it does not follow that physics says *all there is to say* about the fundamental level. The fundamental may be thought to include some logical facts, but it does not follow that physics must incorporate the study of logic. More to our present point, if I am right about the logic of potentiality, then any way that things are fundamentally entails their having the potentiality to be that way, and I see no reason why that potentiality should not equally count as fundamental (see section 5 for more on this). If, further, some of the way things are are not themselves conditional (if, for instance, some fundamental properties are categorical), then we will have nonconditional fundamental potentialities. This is not enough for present purposes since I have not yet argued for any potentialities that might ground non-actualized, non-conditional possibilities. Still, these considerations open up some space for rejecting assumption 2 and evading Bird's problem.

Alternatively, we might question assumption 3: Are all nomological dispositions conditional in form? In some other cases of apparently conditional dispositions, such as water-solubility, I have argued that the manifestation is really to be understood in *causal* terms (Vetter 2015: 96-98; Vetter 2014: 148-151). Thus water-solubility, on my view, is not the disposition to dissolve *if* put in water, but rather the disposition to dissolve-in-water, that is, to be *caused* by (immersion in) water to dissolve. Perhaps I was wrong to build conditionals into the manifestation of the nomological dispositions; perhaps something like this causal story would work better there too.[3] A disposition to be caused by Φing to Ψ seems a better candidate for implying the possibility of both Φing and Ψing than a disposition to Ψ-if-Φ. But it remains to be seen, first, how this is best integrated with the quantitative nature of the nomological dispositions, and second, whether it is borne out by our best (philosophy of) physics.

There are, then, some directions in which one might go to solve Bird's problem of non-conditional possibilities. But I suspect that the problem will trouble dispositionalists for some time to come.

2. *Degrees of potentiality and possible worlds: response to Wildman*

Nathan Wildman offers three objections to the account of potentiality and of possibility that I develop in Vetter 2015. I will here take up the first objection.[4]

Wildman's objection (in section 2 of his paper) is directed against the proportional understanding of the graded possibility involved in dispositions which I have set out in section 1 in response to Bird's criticism. To have a disposition, on the view I defend, is simply to have a potentiality to a sufficient degree; and degrees are best captured in comparative terms. Wildman captures my view of such comparative degrees in the principle

> PROPORTION *x* has [a] potentiality [to] *P* to a greater degree than *y* iff the proportion of worlds where *x* has its relevant intrinsic features and *P* s is greater than the proportion of worlds where *y* has its relevant intrinsic features and *P* s. (Wildman: 172)

[3] For independent reasons, I do believe I was wrong to build a material conditional into the manifestation of nomological dispositions. For as Ralf Busse has pointed out, that threatens triviality: things might possess the disposition to Φ-if-Ψ simply by having the disposition to not-Ψ (see Busse 2015). Whatever solves this problem might solve the problem of non-conditional possibilities as well.

[4] I suspect that the second objection (appealing to the distinction between being destructible and being perishable) can be dealt with on the level of semantics, not metaphysics, by writing the relevant kind membership as an additional condition into the truth conditions; or perhaps even on the level of pragmatics. The third objection can be avoided, as Wildman points out, by adopting 'permanentism' or eternalism. For reasons sketched in Vetter 2015, ch. 7.9, I believe that I am committed to that view anyway.

That is, indeed, the gist of my conception of degrees, insofar as it is to be captured in terms of possible worlds – and insofar as it applies to intrinsic potentialities. With extrinsic potentialities, we would have to add the relevant extrinsic features along with the intrinsic ones.

Wildman then develops an ingenious counterexample to PROPORTION, which has the following structure. Let *P* be a process with a certain necessary precondition (in Wildman's example, *P* is killing humans by venom, and the necessary precondition is the existence of humans). Now let *y*, but not *x*, be essentially tied to the obtaining of that precondition (in Wildman's example, *y* is a cybernetic cobra that is essentially made by humans, *x* is an ordinary cobra with no such essential ties to humans). Then the relevant worlds for *x* will include a large class of worlds where *x* does not P simply because the precondition is not met. Since *y* is by its essence precluded from existing in such worlds, the relevant worlds for *y* will not include a corresponding class. As a result, the proportions are skewed: *y* (the cybernetic cobra) will have a much higher proportion of *P* -worlds among the relevant worlds than *x* even if *x* intuitively has the potentiality to *P* to a greater degree (i.e., has a higher proportion of *P* -worlds among those where the necessary condition is satisfied).

So we have a counterexample to PROPORTION: *x* has the potentiality to *P* to a greater degree than *y*, but its proportion of *P* -worlds among the relevant worlds is not higher, and may indeed be much lower, than *y*'s. Wildman concludes that the "possibility of such cases strongly calls into question understanding talk of potentiality degrees in terms of talk of proportions of worlds" (Wildman: 175).

I would like to consider three possible responses to this counterexample.

A first response is to point out that Wildman's example is one of an *extrinsic*, not an intrinsic potentiality. Whether or not something is venomous to humans plausibly depends on humans, in two ways: it depends on the existence of humans, and on their physiology. But for extrinsic potentialities, PROPORTION cannot be upheld as it stands in any case: we will have to consider worlds where not only *x* has its relevant intrinsic features, but where the relevant external factors also hold. Worlds in which humans have evolved to have a different physiology that makes them immune to the cobra's venome are not relevant; and neither are worlds where there are no humans. Writing the existence and physiology of humans into the conditions that circumscribe the relevant worlds clearly gets rid of Wildman's counterexample: we need not consider those worlds in which *x*, the ordinary cobra, exists unaccompanied by any humans. However, this response relies on a specific feature of Wildman's case: its extrinsicality. It may not apply to other similar cases; but perhaps it can be generalized.

A second strategy thus generalizes the first. The first response disposed of the troublesome worlds (those where *x* exists but the preconditions for *P* ing

are not met) by requiring that certain extrinsic features of the objects are held fixed. But we could dispose of them more directly, by simply requiring that, in addition to keeping fixed the relevant intrinsic features of the object in question, at the relevant worlds all relevant preconditions for their *P* ing are met. (It is a difficult question what makes a precondition 'relevant'; but the same is true for the question what makes an intrinsic feature 'relevant', so I think we can justify postponing that question.) Since the existence of humans is a precondition for their being poisoned, we have again excluded the troublesome worlds in Wildman's example, and have done so in a way that generalizes beyond the extrinsic potentialities. It is obvious that such a response will need to spell out the notion of a 'precondition'. It might do so in conceptual or logical terms; but it might also do so in genuinely modal terms: C is a precondition for *P* ing iff, necessarily, if anything *P* s then C. Would it be circular for the potentiality theorist to appeal to modality at this stage, in understanding potentiality itself? No, since PROPORTION is at any rate not meant as a reductive account of degrees. It is merely a formal model, used to capture the formal structure of potentialities' degrees.

This brings me to a third, and indeed my preferred response, which may be combined with the previous one but can also stand on its own. It is that PROPORTION was never meant to be more than a formal model, capturing or at least approximating the formal structure of degrees of potentiality. Wildman considers a response along these lines and complains that it "makes potentiality degrees even more mysterious" (Wildman: 176) and that "for those of us who struggle to understand potentiality degrees, this [kind of response] is cold comfort" (*ibid*.). I respond that Wildman does not seem to struggle to understand potentiality degrees: he has a very clear grip, in his own counterexample to PROPORTION, on the question of which cobra is more venomous than the other. That kind of grip, like our knowledge of what can and can't happen, is prior to any possible-worlds semantics; the semantics, after all, is modelled in such a way as to capture our intuitive judgements. Wildman's point, of course, is that he and others fail to have a more theoretical, philosophical understanding of potentiality degrees except in terms of possible worlds. I agree that more needs to be said about potentiality degrees than I do in my book (I have tried to say a little more in Vetter 2018b; see also section 3). What needs to be done, I think, is to carefully reflect on our pretheoretic judgements of potentiality degrees (of this being more fragile than that, of one person being more able to run than another) and their inferential relationships, and to formulate general principles about them (such as a transitivity principle, or a complementarity principle for which I have argued: the more *x* is disposed to *P* , the less *x* is disposed to non-*P* , and vice versa), which can then be tested against further reflections of our pretheoretic

judgement – and so on, until we have reached a kind of equilibrium. What more can we do? The theory I propose takes potentiality as its primitive. New primitives are often met with complaints of unintelligibility. The best that their proponents can do is work with them and show them to be fruitful.

3. *Dispositionalist necessity and the role of causation: response to Yates*

In an earlier paper (Yates 2015), David Yates argued that dispositionalism in its most straightforward form (the form I defend) fails to be formally adequate since it does not provide the dispositions we need to distinguish between necessary truths (such as, 2+2=4) and necessary falsehoods (such as, 2+2=5). Responding to his paper gave me the opportunity to elaborate further on how I understand degrees of dispositions or potentiality (see Vetter 2018b).[5]

Degrees, I argued, are best understood so as to give rise to a principle of proportionality, such that the degree of any object *x*'s potentiality to Φ is always indirectly proportional to the degree of *x*'s potentiality not to Φ. This principle of proportionality, in turn, gives rise to what Yates (in this issue) calls universality: the claim that for all *x* and Φ, it is always true either that *x* is disposed to Φ or that *x* is disposed not to Φ (or, of course, both). Universality in turn implies that objects do have dispositions which are always necessarily manifesting, such as a disposition to be dancing-or-not-dancing, or indeed a disposition to be such that 2+2=4. Yates had briefly considered but swiftly rejected this response in his 2015 paper under the title "a plenitude of powers".

The ultimate source of our disagreement, though, is not a matter of whether ontology ought to be plenitudinous or sparse. It is, as Yates makes very clear in his contribution to this issue, our different take on the relation between dispositionality and causation. Yates takes dispositions to be essentially linked with causation; even if we went beyond efficient causal powers (i.e., dispositions to cause events) in our theory of dispositions, we must still maintain some link between dispositions and causation. I, on the other hand, stress the modal aspect of dispositions: they concern what *can* be, and while that is often (and especially in the cases that interest us) linked with causation, it need not be so.

Yates's argument proceeds, not by tackling this disagreement headon, but by arguing against my principles of proportionality and universality. He agrees that the former implies the latter, and provides counterexamples against both. The counterexamples initially rely on the standard model of dispositions as

[5] Since Yates prefers the term 'disposition' to my 'potentiality', and I did the same in my response (Vetter 2018b) to his earlier paper (Yates 2015), I will in this section use two terms interchangeably. See section 4 for a more precise explanation of how I intended to distinguish them in Vetter 2015.

coming with a stimulus and a separate manifestation. The disposition to sing when it rains, for instance, does not appear to be indirectly proportional in its degree to the disposition not to sing when it rains; in fact, Yates argues, one might lack both because rain simply makes no difference to whether or not one sings.[6]

Now, I do not accept the standard model in which a disposition is characterized by a stimulus and a manifestation. Yates notes that "it is not clear […] whether or not [I take] the argument from degrees to depend on [my alternative conception of dispositions]" (105 fn 13). That is because I had hoped that the argument did not so depend. But Yates's argument convinces me that it does. So let me outline how my manifestation-only conception of dispositions accommodates Yates's apparent counterexamples.

As Yates points out, my way of dealing with such apparently stimulus-involving dispositions as the disposition to sing when it rains is to pack it all into the manifestation: what Yates has in mind, on my view, is the potentiality to be-caused-by-rain-to-sing. Applying the principle of proportionality to this potentiality, Yates asks where we should apply the negation: if we give the negation narrow scope, then the potentiality to be-caused-by-rain-to-sing must be indirectly proportional to the potentiality to be-caused-by-rain-*not*-to-sing; if we give the negation wide scope, then our potentiality must be indirectly proportional to the potentiality *not*-to-be-caused-by-rain-to-sing. Yates favours the narrow-scope view and argues, convincingly, that it does not yield proportionality.

However, it should be clear that this is not an instance of the principle of proportionality as I have stated it. On my view, a potentiality comes with a manifestation only. The proportionality principle applies to potentialities with contradictory manifestations: the potentiality to Φ, and the potentiality not to Φ. If we replace Φ with "be caused by rain to sing", then the potentiality not to Φ is the potentiality not to be caused by rain to sing (i.e., it uses the wide-scope negation of Φ). Any temptation to go for a narrow-scope negation is driven by the idea that it is singing, and not the complex property of being-caused-by-rain-to-sing, which is somehow the 'real' manifestation of the potentiality in question. But my view leaves no room for a 'real' manifestation *within* the manifestation.

Given the wide-scope reading of the negation, Yates agrees that proportionality and hence universality succeeds. In fact, he claims that on this reading "universality would be a logical truth: either *x* is disposed to Φ, or not [*x* is dis-

[6] I might point out here that I accept dispositions, or rather potentialities, of very low degrees, so the probability of one's singing given rain would not have to be high to warrant the ascription of a potentiality to sing when it rains. This takes care of the final remark. But it does not make the proportionality principle any more appealing when applied to Yates's case.

posed to Φ]" (Yates: 107), because "the most natural way to render '*x* is disposed not to be caused to sing by rain' in the stimulus-manifestation idiom is: not [*x* is disposed to sing when it rains]" (Yates: 106). Now, what is or isn't the most natural reading of a sentence in the stimulus-manifestation idiom is neither here nor there, since we are dealing with the alternative, manifestation-only, conception.[7] But there is an independent objection in the remark that I have quoted: it says that the wide-scope reading collapses into a negated disposition ascription. That would make my principle trivial indeed, but it is not true. We can distinguish between something's having a potentiality not to be caused by rain to sing, and its lacking the potentiality to be caused by rain to sing. The former, but not the latter, comes in degrees: one can be more or less disposed not to be caused by rain to sing, with the maximal degree amounting to a (relative) necessity of never being caused by rain to sing. Of course, the distinction is thin, given my own argument for proportionality, but it is nevertheless there.

I do, therefore, hold on to the principles of proportionality and universality, and continue to hold that there are potentialities which are necessarily always manifested, such as the potentiality to be such that 2+2=4. Yates, however, has another objection. Even if proportionality and universality hold for values of Φ that are suitable as manifestations of dispositions, the argument should not be carried over to such properties as *being such that 2+2=4*, for that property is simply not embedded in the causal nexus. It is here that we return to our ultimate disagreement: how do dispositions relate to causation? In Vetter 2018b, I cite examples from Nolan 2015, as well as from physics to show that some dispositions have manifestations that are non-causal. However, Yates points out that Nolan's dispositions can at least be taken to be grounded in more fundamental causal dispositions, and that it is controversial whether the examples from physics are really non-causal. Neither applies to the disposition or potentiality to be such that 2+2=4, if there is one: it is uncontroversially and fundamentally uncausal.

We have come full circle to our initial and basic disagreement: are there dispositions that are entirely acausal, i.e., not embedded at all in the causal nexus? Yates insists that there aren't, while I hold that there are. I do agree with Yates that our initial and paradigmatic examples of dispositions and hence of potentialities, both from ordinary life and from science, are causal in some sense. If they were not, then recognizing them would be less useful for our practical purposes of manipulating objects and predicting their behaviour.

[7] Yates may have in mind here the idea that the stimulus-manifestation idiom is so much more natural that I must be able to translate my own idiom back into it. I tend to believe, and I have argued in Vetter 2014, that the stimulus-manifestation idiom has no pretheoretical force, and I would venture to claim that appearances to the contrary stem from contemporary philosophers being raised on a dogma that dates back to classical empiricism and its concern with verification.

In philosophy, we inevitably generalize beyond the initial and paradigmatic examples. In doing so in the present case, we *can* hold on to the causal element and stop where it gives out; we may even reserve a term, be it "power" or even "disposition", for the result of such a generalization. But note that this would be a decision, not an analysis of an already established usage: "disposition", in the philosopher's vernacular, is a theoretical term. And instead of stopping where causation gives out, we can also generalize further and notice that there is a modal element involved in our initial range of cases which can in principle be separated from the causal one. It is certainly not a conceptual confusion to say that some fundamental properties are dispositional (in the sense of "are like our initial examples of dispositions") but acausal; Yates provides reasons against the truth of such a claim, but not against its conceptual coherence. (For more examples, see Vetter 2015: 98.) Thus it seems we *can* generalize beyond the confines of causal dispositions; and if we can, why shouldn't we do so where it fits our theoretical purposes? One response to this question is that without integration into the causal network we have no reason to believe that such properties really exist. But I would beg to differ: causation is but one kind of explanation, and we might as well take integration into the *explanatory* order of the world as evidence of existence. (In fact, I am inclined to think that the big metaphysical questions are not questions about existence at all, but are rather questions about the explanatory order of the world; see also below, section 5.) And if that is so, then I believe that there is plenty of space for such properties as maximal potentialities, which explain, for instance, why things have no potential for doing otherwise.

This does not, of course, settle the debate but merely serves to point to the more general and in some sense deeper issues that underly it: David Yates's and my different conceptions of dispositions are rooted, I suggest, in our different conceptions of metaphysics itself. But to discuss those directly is a task that I must leave for another time.

4. *Degrees, dispositions, and the metaphysics of potentiality: response to McKitrick*

Potentiality is rather handwaving about the exact metaphysics of potentialities: are they universals, and if so, are they universals of the Aristotelian or of the Platonic variety? Or could they be fit into a nominalist metaphysics? My hope was that what I said in the book was compatible with a wide array of answers to these questions. Jennifer McKitrick challenges this hopeful attitude.

There is much that I agree with in McKitrick's paper. She is certainly right, for instance, that neither Class Nominalism nor Resemblance Nominalism are

viable options for an account of potentialities along the lines that I envisage, despite my officially non-committed stance on the metaphysics of properties in *Potentiality*.

Rather than going through her many arguments one by one – which, desirable as it would be, is impossible in the limited space I have here – I would like to address one central premise that runs through several of her arguments, and which concerns the relation of a determinable potentiality to its determinates, the specific degrees of that potentiality. On my behalf, McKitrick reconstructs an argument that appears to show that the determinable potentiality is, on my view, more fundamental than the determinate, degreed, potentialities. She takes this claim not as in itself a *reductio ad absurdum* of my view, but rather as a view which a metaphysics of potentialities must, and which most candidate metaphysics fail to, accommodate.

I do not endorse the claim that a determinable potentiality is more fundamental than its determinate degrees, and I do not think that I am committed to the claim. To show why, let me begin by reproducing the crucial premises of the argument which McKitrick has constructed on my behalf (the premises are quoted from McKitrick: 144; quotations and page numbers within the premises refer to Vetter 2015):

> 3. A disposition is a degree of a potentiality: "having a disposition such as fragility is a matter of having the right potentiality (in this case the potentiality to break or be broken) to a contextually sufficient degree" (22).
> 5. Potentialities ground dispositions: "The notion of a potentiality has been introduced as the metaphysical background to the context-dependent notion of a disposition" (96).

If these two premises are granted, it follows that determinable potentialities ground at least some of their determinates; and then it would be arbitrary not to claim that they ground *all* of their determinates.

I do not accept either premise 3 or premise 5 as stated in McKitrick's words. Explaining why not gives me a welcome opportunity to further clarify the relation between dispositions and potentialities, or rather: between the notion of a disposition and that of a potentiality. In Vetter 2015: 80-84, I claim that the relation between a given disposition term, say 'fragile', and the corresponding potentiality, i.e., the potentiality to break, is analogous to that between 'tall' and height. To clarify why I reject premises 3 and 5, I will again refer to this analogy.

Consider, first, premise 3, and its analogue with 'fragile' replaced by 'tall' and 'potentiality' by 'height':

> 3'. Tallness is a degree of height: "[being tall] is a matter of having [height] to a contextually sufficient degree" (22).

Clearly, 3' is false: tallness is not itself a degree of height. To start with, there is no one property of tallness: 'tall' expresses different properties in different contexts. Given a particular context, of course, 'tall' does express a particular property. But that property is not identical with any particular (degree/determinate of) height. There are many determinate heights that can make an individual satisfy the predicate 'is tall', as interpreted in a given context. We can think of the property expressed by 'is tall' in a given context as involving something like existential quantification: as the property of having *some* determinate height above a given threshold. Thus an individual's satisfying 'is tall', as interpreted in a given context, depends on or is grounded in the individual's particular determinate height property; but the property expressed is not identical with any particular determinate height property.

For exactly analogous reasons, 3 is false: fragility is not itself a degree of the potentiality to break. To start with, there is no one property of fragility: 'fragile' expresses different properties in different contexts.Given a particular context, of course, 'fragile' does express a particular property. But that property is not identical with any given degree of the potentiality to break. There are many determinate degrees of the potentiality to break that can make an object satisfy the predicate 'is fragile', as interpreted in a given context. We can think of the property expressed as involving something like existential quantification: as the property of having *some* determinate degree of the potentiality to break that is above a given threshold. Thus an individual's satisfying 'is fragile', as interpreted in a given context, depends on or is grounded in the individual's determinate of the potentiality to break; but the property expressed is not identical with any particular determinate potentiality.

Thus premise 3 is misleading: it is not true that a disposition *is* a degree of a potentiality, though it is true that the property expressed by a dispositional predicate in a given context depends on the degree of the potentiality.

Premise 5, too, is false, as is its analogue with 'tall' and 'height:

> 5'. Height ground tallness: "The notion of [height] has been introduced as the metaphysical background to the context-dependent notion of [tallness]" (96).[8]

[8] 5' is false also because that is not how the notion of height has been introduced into discourse. But let's disregard this disanalogy; it does nothing to undermine the analogy that I am after.

With the term 'metaphysical background', as quoted in 5, I did not intend to introduce yet another redescription of grounding. Rather, I meant to describe the contrast between a *semantic* phenomenon, the context-sensitive variability of a predicate's intension, and the metaphysics that provides the material for that variation. The metaphysical background in this sense is the range of phenomena from which the context-sensitive expression picks its semantic values, depending on the context.

Height, both the determinable and its determinates, are the metaphysical background for 'tall', since it is from these properties that any context selects semantic values for 'tall' (not by picking one, but by setting a threshold, as described above). This is not to say, as 5' says, that height (the determinable) grounds tallness. Rather, and as we have seen above, it is the instantiated determinate that does the grounding in any particular case.

Likewise, potentiality, both determinable and determinate, is the metaphysical background for 'fragile', since it is from these properties that any context selects semantic values for 'fragile' (not by picking one, but by setting a threshold, as described above). This is not to say, as 5 says, that the determinable potentiality to break grounds fragility. Rather, and as we have seen above, it is the instantiated determinate/degree that does the grounding in any particular case.

In short, 'potentiality' is contrasted with 'disposition' not as determinable with determinate, but rather as the metaphysical level, including both determinable and determinates, with the semantic. Once this is recognized, it should be clear that the argument does not go through, and we have no more reason to accept McKitrick's conclusion:

> 8. Therefore, determinable potentialities are more fundamental than their determinate dispositions: "the general dispositions are not only equally fundamental as the specific ones, they are more fundamental" (57),

than we do to accept the analogous

> 8' Therefore, height is more fundamental than its individual determinates.

But if I reject McKitrick's argument, why do I seem to endorse its conclusion in the quotation she gives under 8?

The quotation comes from a passage where I discuss, and ultimately reject, the conditional conception of dispositions (see section 1). Within that conception, I argue, we must distinguish between the general disposition to break if struck, or to exert a force of $F = \in(eq/r^2)$ when at distance r from an object with

charge q; and the specific disposition to break if struck with a force of 8.35N, or to exert a repulsive force of $8\times10^{-8}N$ when at a distance of $5.3\times10^{-11}m$ from a charge of $1.6\times10^{-19}C$. With that distinction in place, I ask which of these dispositions are more fundamental, the general or the specific. I argue that the conditional conception favours the specific dispositions, while independent philosophical considerations on grounding and fundamentality would favour the general dispositions, thus questioning the adequacy of the conditional conception. The independent philosophical considerations that I adduce draw strongly on Jessica Wilson's arguments to the effect that determinables can be fundamental. But I am explicit that general dispositions do not relate to specific ones as determinables to determinates (Vetter 2015: 53, 55): unlike a determinable, whose instantiation necessitates the instantiation of one of its determinates to the exclusion of all others, instantiating a general disposition necessitates instantiation of all or at least very many of the corresponding specific dispositions.

While my argument about general and specific dispositions thus makes use of Wilson's arguments, I never make the corresponding claim about determinable potentialities and their degree-determinates, let alone the stronger claim expressed in McKitrick's 8. My reasons for not making that claim are very much the same reasons that McKitrick gives on p. 15. Exactly how we are to understand the relation between the determinable potentiality and its determinates, the individual degrees of potentiality, is a difficult question that I do not address in the book. I suspect, however, that it will be a question that is not specific to the metaphysics of potentiality but rather to be answered by general considerations about determinables and determinates. McKitrick may very well be right that an answer to this question will have repercussions for our ontology of potentialities in particular, and of properties in general.

The main question posed by McKitrick remains, of course: what is the best metaphysics for the properties that I call potentialities? I will take this question up again in the next section.

5. *The metaphysics of potentiality, grounding, and counterpossibles: reponse to Giannini and Tugby*

Like Jennifer McKitrick, Giacomo Giannini and Matthew Tugby ask about the metaphysics of potentialities: are they universals, and if so, are they best understood in Platonist or in Aristotelian terms? Their arguments, which I take to be offered in a constructive spirit, suggest that we (or, at any, rate, I) had better think of potentialities as Platonic universals.

Tugby (2013) has forcefully argued that dispositionalists are committed to a Platonist conception of properties; very roughly, the argument is that if a dis-

position is individuated (at least in part) by its manifestation, then if anything is to have a disposition to Φ, there must be some property of Φing in the first place. On both nominalist and Aristotelian theories of properties, the existence of such a property is dependent on its being instantiated, sometimes and somewhere; only on a Platonist conception could there be a property of Φing even if nothing has ever, and nothing ever will, Φ. But clearly things could have such dispositions; hence we should adopt the Platonist conception.

My resistance to such a full-blown Platonist conception is attributed, by Giannini and Tugby, to 'ontological naturalism', the "doctrine that reality consists of nothing but a single all-embracing spatio-temporal system" (Armstrong 1981: 149, cited on page 75). I am not sure that I am an ontological naturalist in this sense; I have certainly not excluded the existence of abstract objects (see Vetter 2015: ch.7.7). As I have briefly indicated above (section 3), I tend to think that the big metaphysical questions are, *pace* the Quinean tradition, not questions about ontology in the sense of 'what there is', but are rather questions about the explanatory order of the world or 'what grounds what' in Schaffer (2009)'s useful turn of phrase.[9] What does motivate my approach is not so much a restriction of what there is to the spatio-temporal, but rather an Aristotelian commitment that Wang, in her contribution to this issue, captures with the term '*de re* first'. It is simply the idea that our world is primarily one of objects, and that objects should be given pride of place in our metaphysical theories. Among the objects in our world, *concrete*, spatiotemporal objects are certainly paradigmatic, and one motivation for the theory is that we can thus start with something that is deeply familiar and epistemically accessible. But clearly my theory goes far beyond the familiar and ordinarily accessible, and I do countenance abstract objects. So the focus on concrete, ordinary objects is not so much a matter of imposing the restrictive claim that is made by ontological naturalism, but rather the implementation of another Aristotelian idea, that we should start with the familiar (even if the ultimate shape of our theory will turn out rather unfamiliar).

For these reasons, I would not describe myself as an ontological naturalist in the sense at issue in Giannini and Tugby's paper. Nevertheless, I prefer an Aristotelian view of properties to full-blown Platonism precisely because Aristotelianism puts objects first. In Vetter 2015, ch.7.5, I suggest that the Aristotelian approach is right in thinking that which properties exist depends on how objects are. But unlike standard Aristotelian approaches, I do not claim that a property exists iff it is (sometimes, somewhere) instantiated. Instead,

[9] For more on this, see Vetter 2018a and Vetter ms.b; in Vetter 2015, this tendency is manifested in the ontological liberalism professed on p. 29.

I suggest that a property exists iff it is instantiated, or potentially instantiated, or potentially potentially instantiated, and so forth – in short, there is a property of being Φ just in case something has (or some things have) an iterated potentiality for something to be Φ. (Given axiom T for potentiality, this includes the case where something actually is Φ.) In this way, we obtain a great many more properties than those which happen to be instantiated, and ensure that every potentiality has a property to serve as its manifestation. But unlike the Platonist, we do not make those properties quite independent of the instantiating objects. Giannini and Tugby put this by saying that on my view even "unmanifested properties exist in the sense that they are *grounded* in the potentialities of *things*" (129, second emphasis mine; I will note my reservations about the formulation below).

Giannini and Tugby note that this view is "dangerously (or, [they] think, fortunately) close to Platonism" (p. 14), and I agree: the main difference is that on my view properties must still be ultimately grounded in objects, thus satisfying my *de re* first approach. Against this view (and hence in favour of full-blown Platonism), Giannini and Tugby object that it "seems incoherent to suppose that potentialities, which are ontologically fundamental, could be individuated by something less fundamental than themselves and which they themselves ground" (130). I see two related objections here. The first is explicit and concerns a 'principle of purity', which says that potentialities, being ontologically fundamental, could not be individuated by the less fundamental properties that are their uninstantiated manifestation properties. The second, which is only implicit, is a circularity problem: how can the potentiality provide the ground for the manifestation, when it is itself (via its individuation) grounded in the manifestation property?

To respond to both objections, I would like to clarify the grounding picture that is suggested in my Aristotelian view of properties. The Aristotelian picture I suggest is one on which objects, by being some way or another, ground their properties. But they do not ground them one by one. Rather, on the dispositionalist picture, we can think of properties as nodes in a vast network held together by the manifestation relation. By instantiating any one property, an object gives reality to the whole network to which it belongs. Despite the metaphorical nature of the description, I hope it is clear how it disarms the circularity worry: by denying that a potentiality grounds its manifestation. Rather, it is objects that ground both the potentiality and the manifestation, but by instantiating only one of them, the potentiality. So I do not, after all, want to say with Giannini and Tugby that "unmanifested properties exist in the sense that they are *grounded* in the potentialities of things"; rather, I want to say that unmanifested properties exist because they

are grounded in things, and more specifically in those things which have a potentiality for their instantiation.

What about purity? Since potentialities do not, as I have just argued, ground their manifestations, the manifestation need not be less fundamental than the potentiality after all, and we can individuate a potentiality in terms of its manifestation without violating a principle of purity.

We might worry that potentialities aren't fundamental after all if they, along with their manifestation properties, are grounded in objects. But in saying that (some) potentialities are fundamental, I never intended to contrast them with entities of other categories (although I did not make that explicit). The contrast, after all, is with Humeanism. According to Humeans, the properties at the fundamental level of nature are all categorical; according to dispositionalism, the properties at the fundamental level of nature, if there is one,[10] at least include dispositions or potentialities. The fundamentality claim is applied within the realm of properties, and should be independent of whether there is a further dependence relation between properties quite generally and objects.

So far, I have given some motivation for the Aristotelian view of properties that I suggest in Vetter 2015, ch. 7.5, and I have tried to defend it against an objection. But for all I have said, Platonism might still be the better metaphysics of potentialities, despite going against my *de re* first approach.

One intriguing reason for adopting Platonism is given in section 3 of Giannini and Tugby's paper: only Platonism allows for superalien properties (properties for whose instantiation nothing has even an iterated potentiality), which in turn are needed to make (dispositionalist) sense of certain scientific claims, to wit, counternomic or counterlegal conditionals in the context of idealizations (cf. p. 131-137). Giannini and Tugby note that both a fictionalist approach and my own view that counterpossibles often concern epistemic rather than metaphysical modality would provide a solution here, but argue that both solutions "incur the cost of leaving us with a disunified treatment of scientific modal discourse" (137). As mere pointers toward possible responses, I want to mention two things. First, how unified the relevant scientific practices are is an empirical matter which it is difficult to judge from the armchair; we would need philosophically informed sociology of science to establish it. Second, even if there is a unified practice, this does not entail that unified truth conditions underly it. As Emanuel Viebahn and I have argued elsewhere for the case of modal auxiliaries (Viebahn and Vetter 2016), different kinds of facts may play the same role

[10] I am now inclined to think that the debate should not be framed as one about what there is at the fundamental level, but rather as one about explanatory hierarchies which may or may not terminate at a fundamental level: see, again, Vetter 2018a and Vetter ms.b.

in our practices and for that reason be expressed with the same kinds of terms; we should not then infer from sameness of expressions to sameness of truth-conditions. This said, it is obvious that the question of super-aliens and counter-nomics is one that certainly merits more detailed consideration than I can give it here, and one on which dispositionalists can and will reasonably disagree.

6. *Metaphysical modality, time, and methods: response to Wang*

Jennifer Wang characterizes my theory as "*de re* first" and contrasts it with her own, incompatibility-based view of modality, which is "*de dicto* first". I wholeheartedly endorse the characterization of my approach as *de re* first: giving pride of place to *objects* in our metaphysics is indeed one of the foremost motivations for the approach. Wang raises two objections against my particular version of a *de re* first account.

Wang's first objection arises directly from my focus on objects. It is the worry that the approach cannot account for paradigmatically *de dicto* modal truths: necessarily, all squares have linear sides; necessarily all yellow things are coloured (both on p.6); and the sentence numbered on page 193:

> (4) Necessarily, no negatively charged objects are positively charged.

On a *de re* first approach, Wang suggests, we cannot get the required *de dicto* readings of truths like (4). All we get are truths such as "[a]ll negatively charged objects are necessarily not positively charged. But", Wang objects, "this is still only a *de re* modal predication rather than a *de dicto* claim" (193).

Let us be clear what the challenge is. My task is to provide, in terms of my potentiality-based theory, a *de dicto* reading of sentence (4): a reading, that is, on which the necessity operator takes scope over the quantifier 'no' in (4).

But this is a requirement that my approach can satisfy. The necessity operator, on my view, is defined as follows: it is necessary that *p* just in case nothing has, had, or will have an iterated potentiality for non-*p*. Plugging that into the *de dicto* statement (4), we get:

> (4DD) Nothing has, had, or will have an iterated potentiality for it not to be the case that no negatively charged object that is positively charged.
>
> (Or, eliminating the double negation: Nothing has, had, or will have an iterated potentiality for there to be a negatively charged object that is positively charged.)

By contrast, the *de re* statement given by Wang will read:

> (4DR) All negatively charged objects are such that nothing has, had or will have an iterated potentiality for *them* to be positively charged.

(4DD) is '*de re*' in the sense that it quantifies over objects first; it has a quantifier that takes wide scope (i.e. it scopes over the potentiality operator, the negation, and the other quantifier). This is because my account sees a hidden quantifier 'within' the necessity operator. But my task was not to get rid of wide-scope quantification over objects altogether; it was to make sure that the explicit quantifier in (4) has narrow scope; and that it does, very clearly, in (4DD). (4DD), in combination with the potentiality-based theory of modality, does just what a *de dicto* reading of (4) should do: it excludes that there are potentialities, hence possibilities, for anything to be both negatively and positively charged; it does not exclude potentialities, and hence possibilities, for any of the actually negatively charged objects to change and become positively charged instead. (And as it should, (4DR) does the exact opposite.)

In short, the potentiality-based approach is '*de re* first' in the sense that it reduces modality to how *things* are. It does not follow that the approach cannot allow for *de dicto* readings of modal *sentences*.

Wang's second objection is also connected to my approach's focus on objects, but with a more specific twist. Objects are typically contingent and temporal entities. They come into and go out of existence, and while existing they change in various ways. My account appeals to just such changing features of individuals: their potentialities. Metaphysical modality, however, is supposed to be non-contingent and atemporal. This is the tension that Wang's paper makes very clear. Her preferred account, based on incompatibility relations between properties, does not face the same problem. Properties (unlike their instantiation by objects) do not appear to exhibit the contingency and temporality that objects do.

Let me first point out that there are various ways to go here. One way to go is to remain *dispositionalist* but cease to be *de re* first, by quantifying, in the definition of modality, over properties and not objects. (Roughly: It is possible that *p* just in case *there is a property* Φ which is a power to bring it about that *p*.) That strategy is endorsed by some (Jacobs 2010, Yates 2015; I believe that Giannini and Tugby would be sympathetic as well) and, depending on the view of properties it is paired with, holds some promise of overcoming the contingency and temporality associated with objects. A second way is to remain *de re* first but cease to be *dispositionalist*, by adopting instead an essentialist account. Truths of essence, while still about objects, have been argued to be not just

atemporal and non-contingent, but in a certain sense even outside the temporal and modal realm (Fine 2005).[11] Still, I would like to defend the combined dispositionalist and *de re first* account; so I need to respond to Wang's worries.

More specifically, Wang's second objection concerns potentiality's relation with time. As she points out, the account seems to be unable to accommodate possibility claims such as her sentence (194)

> (5) It's possible for there to be an object that always exists in a universe with no beginning.

Wang says that she considers such cases not as counterexamples to my view, but as "unintuitive consequence[s]" (195) that "undermine a crucial selling point of [my] theory: its intuitive attractiveness" (6).

How is one to argue about intuitions, or alleged intuitions, such as (5)? When debating metaphysical modality, philosophers often take it to be unproblematic that we are all talking about the same thing and genuinely disagreeing about it. But there is no pretheoretical, philosophically neutral concept of, nor is there unproblematic reference to, metaphysical modality that we can use to focus on the phenomenon prior to giving a particular theory of it: unlike, say, knowledge, metaphysical modality is a theoretical concept.

There are, of course, related phenomena that it might be easier to get a grip on. One is our ordinary, everyday understanding of modality that qualifies as objective, albeit not as metaphysical: I can ride my bike to work, but I cannot get from Berlin to Milano in less than an hour. The other related phenomenon is logical modality, on which we have perhaps a firmer, or at any rate a formally regimented, theoretical grip. Metaphysical modality is uncomfortably wedged between the two. In our thinking about genuinely metaphysical modality, we might either start from the logical notion and see how we can account for the cases where logical does not entail metaphysical possibility; this is a route taken, in very different ways, e.g. in Hale 2013 and Chalmers 2010. Or else we might start with the ordinary notion and see how we can account for the cases where ordinary impossibility does not entail metaphysical impossibility; this is a route taken or suggested, again in different ways, in Williamson 2007 and Edgington 2004. The second route is clearly the more congenial to a potentiality-based approach like mine, which starts from our ordinary understanding of what we and the objects around us can and cannot do. Wang, on the other hand, seems to incline towards the first approach, e.g. when she writes, about the case of a glass appearing *ex nihilo*, that "[t]here is no logical impossibility,

[11] This, incidentally, is one of the reasons why *pace* Wang (p.4), essence and potentiality are not duals; see Vetter ms.a.

and hence, according to many, no metaphysical impossibility involved in the supposition" (194).

Does it matter where we start in talking about metaphysical modality? I am inclined to think that it does. Starting with the wider notion of logical possibility, we introduce restrictions: such-and-such is logically possible, but it is not metaphysically possible because… The default for a proposition, at least if it is logically consistent, is possibility; it is claims of impossibility that require justification. Starting with the narrower notion that we express with everyday modals, we must instead proceed by extending the scope of our modal concept: such-and-such is impossible given the state of our technology, but it is metaphysically possible because… Possibility is not the default but rather something that needs to be justified and supported. (Much of the argument in Vetter 2015 can be seen as giving this kind of justification, starting from ordinary modality in the form or dispositions and abilities.)

It is unsurprising, then, that these two starting points engender different intuitions. Are they even intuitions about the same phenomenon, or are philosophers talking past each other when coming from these two different starting points? This, I submit, is an open question that should receive more attention than it has so far received. What I would like to claim here is simply that things aren't so clearcut when it comes to intuitions about metaphysical modality. What is intuitive depends on where we start in getting a grip on this theoretical notion. Wang's (5) seems highly intuitive when, like her, we start with logical modality and require positive justification for any necessity that is not logical necessity. It seems, I submit, much less intuitive when, like me, we start with ordinary modality and require positive justification for any possibility that outstrips those possibilities which we countenance in everyday life.

I will end, thus, with a desideratum. In thinking about metaphysical modality, we need to reflect and make explicit where we start; and we need to reflect on whether and why we can assume that we are all targetting a common, albeit theoretically circumscribed, phenomenon: metaphysical modality.

Barbara Vetter
barbara.vetter@fu-berlin.de
Freie Universität, Berlin

References

Aimar, Simona, 2018, "Disposition Ascriptions", in *Philosophical Studies*, https://doi.org/10.1007/s11098-018-1084-9.

Armstrong, David, 1981, *The Nature of Mind and Other Essays*, Cornell University Press, Ithaca.

Bird, Alexander, 2007, *Nature's Metaphysics*, Oxford University Press, Oxford.

Busse, Ralf, 2015, "Zur These der Reduktion von Möglichkeit auf Vermögen", in *Zeitschrift für philosophische Forschung*, 69: 403-407.

Chalmers, David J, 2010, "The Two-Dimensional Argument Against Materialism", in *The Character of Consciousness*, Oxford University Press, Oxford.

Edgington, Dorothy, 2004, "Two Kinds of Possibility", in *Proceedings of the Aristotelian Society*, Supplementary Volume, 78: 1-22.

Fine, Kit, 2005, "Necessity and Non-Existence", in Id., *Modality and Tense. Philosophical Papers*, Oxford University Press, Oxford: 321-354.

Hale, Bob, 2013, *Necessary Beings*, Oxford University Press, Oxford.

Jacobs, Jonathan D, 2010, "A powers theory of modality: or, how I learned to stop worrying and reject possible worlds", in *Philosophical studies*, 151: 227-248.

Lowe, E.J, 2011, "How *Not* to Think of Powers: A Deconstruction of the "Dispositions and Conditionals" Debate", in *The onist*, 94: 19-33.

Nolan, Daniel, 2015, "Noncausal dispositions", in *Noûs*, 49: 425-239.

Schaffer, Jonathan, 2009, "On What Grounds What", in D.J. Chalmers, D. Manley, and R. Wasserman (eds.), *Metametaphysics. New Essays on the Foundation of Ontology*, Oxford University Press, Oxford: 347-383.

Tugby, Matthew, 2013, "Platonic Dispositionalism", in *Mind*, 122: 451-480.

Vetter, Barbara, 2018a, "Evolved Powers, Artefact Powers, and Dispositional Explanations", in *Aristotelian Society*, Supplementary Volume, 92: 277-297.

Vetter, Barbara, 2018b, "A Plenitude of Powers", in *Synthese*, online first: 1-21.

Vetter, Barbara, 2015, *Potentiality: From Dispositions to Modality*, Oxford University Press, Oxford.

Vetter, Barbara, 2014, "Dispositions without Conditionals", in *Mind*, 123: 129-156.

Vetter, Barbara, 2012, "Dispositional Essentialism and the Laws of Nature", in A. Bird, B. Ellis, and H. Sankey (eds.), *Properties, Powers and Structures. Issues in the Metaphysics of Realism*, Routledge, London.

Vetter, Barbara, ms.a., "Essence, Potentiality, and Modality", unpublished manuscript.

Vetter, Barbara, ms.b., "Explanatory Dispositionalism: What the anti-Humean should say", unpublished manuscript.

Viebahn, Emanuel and Vetter, Barbara, 2016, "How many meanings for 'may'? The case for modal polysemy", in *Philosophers' imprint*, 16: 1-26.

Williamson, Timothy, 2007, *The Philosophy of Philosophy*, Blackwell, Oxford.

Yates, David, 2015, "Dispositionalism and the modal operators", in *Philosophy and Phenomenological Research*, 91: 411-424.

Past Present

On madness and free will: a Kantian debate in Denmark in the first half of 19th century

Ingrid Basso

Abstract: The so-called "Howitz-dispute," which arose in Copenhagen in the second half of the 1920s, represents a sort of tear in a monotonous and uniform ideological fabric, whose consequences are destined to last until the middle of the century and ideally join together with the strongest continental currents. The dispute takes its name from the Danish professor of forensic medicine, Frantz Gotthard Howitz (1789-1826), who wrote in 1824 the treatise On Madness and Ascribing Responsibility: A Contribution on Psychology and Jurisprudence, in which he considered a problem to which the entire post-revolutionary civil society was looking for a fair solution, namely, the problematic relationship between madness and the ascription of responsibility. The treatise immediately evoked a number of critical reactions, since the author "accused" the Danish law of the time of being based on Kant's view of morality. Howitz's treatise has the merit of originality not only because, from a chronological point of view, it comes before many of the most important writings on the theme of madness and imputability (and its author is thoroughly acquainted with the international scientific literature on mental illnesses), but also because it shows how at the root of the Danish clinical and legal reasoning of the time, there was the Kantian moral doctrine.

Keywords: determinism; forensic medicine; freedom; Howitz F.G; imputability; Kant, Immanuel; morality; madness; reason; will.

> *Exhaustive, l'identité n'est pas un critère.*
> Michel Foucault (1969: 187)

> *The relative difference here is no sorites from which the quality is supposed to appear by a coup des mains [sudden stroke], since it is within the specific quality.*
> Søren Kierkegaard (1844: 131; *SKS* 4, 288; Eng. tr.: 90)

In a cultural context saturated with German romantic literature and philosophical idealism, the so-called *Howitzfejde*, or "Howitz-dispute", which arose in Copenhagen in the second half of the 1920s, represents a sort of tear in a monotonous and uniform ideological fabric, a "breath of fresh air" (Høffding

ISSN (print) 2281-8618-ETS DOI: 10.4454/philinq.v8i1.283

1909b: 91)[1] destined to last until the middle of the century and ideally join together with the strongest continental currents.

In the wake of the introductory lectures on Schelling's thought, given at Elers Kollegium in 1801 by the Danish-Norwegian philosopher of nature Henrich Steffens (1773-1845) (Steffens 1803),[2] the philosophical Parnassus and in general the Danish intelligentsia was turned in the direction of *Naturphilosophie*. Therefore, the voice of a young Danish professor of forensic medicine, Frantz Gotthard Howitz (1789-1826), who claimed in a deterministic key the concrete reasons for a *homo phenomenon* (to which no *homo noumenon* would be opposed as *imperium in imperio*), was bound to raise a sensation and, above all, be disconcerting.

The real origin of the debate was a problem to which the entire post-revolutionary civil society was looking for a fair solution, namely, the problematic relationship between madness and the ascription of responsibility. In order to answer this question (which, on the juridical and medical side, had in theory already received an almost unanimous answer everywhere on the continent),[3] one had to consider an "epistemological demarcation" of the field of investigation and application: how to qualify madness? This was one of the problems of nascent psychiatric science, and it was this question that the 35-year-old Frantz Gotthard Howitz wanted to address when he wrote his treatise in 1824, *On Madness and Ascribing Responsibility: A Contribution on Psychology and Jurisprudence*.[4] We present here for the first time an English translation of Chapter Seven of this work, which concerns Howitz's criticism of the Kantian doctrine of freedom, by the eminent scholar of the Danish Hegelianism and Danish "Golden Age" Jon Stewart.

Although it appeared later as an independent monograph, the treatise was published initially as an article in the *Juridisk Tidsskrift* (Journal for Jurisprudence), which was directed by the influential and experienced jurist Anders Sandøe Ørsted (1778-1860), the future Danish Prime Minister. Howitz's writing set off the greatest controversy in the history of the Danish philosophy up until that time, namely, the above mentioned *Howitzfejde*. It involved promi-

[1] Cf. Koch (2003c: 357-359). The Danish-German phrenologist Carl Otto (1825: 198) even talks about a "literary slavery", cf. Jacobsen (2007: 65).

[2] See Paul (1973), Koch (2004: 31-56); Stewart (2003: 204ff); Basso (2007: 88-94).

[3] A landmark in this sense was Napoleon's *Code penal* of 1810, see art. 64: "There can be no crime, or delict, where the accused was in a state of madness, at the time of the action, or when he has been constrained by a force which he had not the power to resist" (*The Penal Code of France* 1819: II, 14).

[4] *Om Afsindighed og Tilregnelse, et Bidrag til Psychologien og Retslæren*, in *Juridisk Tidsskrift*, 8, 1, 1824: 1-117. I've edited an Italian translation of the text in 2017, the present introduction is based on my preliminary essay to it. Cf. Basso (2017).

nent figures such as Anders Sandøe Ørsted – brother of the famous physicist Hans Christian (1777-1851) – the theologian and later bishop Jacob Peter Mynster (1775-1854), the aesthetician Johan Ludvig Heiberg (1791-1860),[5] and the professor of philosophy Frederik Christian Sibbern (1785-1872), later mentor of Søren Kierkegaard (1813-1855). It is also possible to find an echo of this debate in Kierkegaard himself.[6] It is worth noting that this dispute would also leave a meaningful legacy if we consider that as late as 1866 the title of the annual Philosophical Prize of the University of Copenhagen was *Did the Controversy on the Actuality of the Human Freedom of the Will that Occurred in the Past in Our Philosophical Literature Leave an Exhaustive and Sound Scientific Result?*[7] (The winner of the essay competition was the philosopher and historian of philosophy Harald Høffding.)

Howitz was a medical examiner, a doctor, named as a member of the Danish College of Health. One of his duties in this capacity was to evaluate the degree of responsibility of criminals, so that the jurists could make some sort of determination about what punishment was appropriate. In 1824 his long article on *Madness and Ascribing Responsibility* appeared. The editor of the journal in which it appeared, Anders Sandøe Ørsted, had written in 1798 a treatise on Kant's theory of freedom, *Over Sammenhængen mellem Dydslæren og Retslærens Princip* (On the Correlation Between the Doctrine of Virtue and the Principle of Law), a book that was characterized by the historian of the Danish philosophy Carl Koch (2003a: 44) as "the most mature fruit of the Danish Kantianism".

Thus, when Howitz's treatise appeared, it immediately evoked a number of critical reactions, since the author "accused" the Danish law of the time of being based on Kant's view of morality. Howitz criticized Kant's conception of freedom as the ability to determine one's own actions based on a correct *rational understanding* of the situation. According to Howitz, it is not true that human beings, as moral agents, are independent of everything empirical. Humans are complex entities, comprised of both elements, the natural and the rational. He argued against Kant's view – according to which there is an *a priori* practical reason that dictates that one follows the moral law – that the moral development essentially depends on the material organization of the brain. We will see in more detail what the objections are that Howitz raised against

[5] On the role of J.L. Heiberg in the dispute see especially Stewart (2007). We do not have here the space for a detailed account of the role of each participants in the dispute, for a complete survey see especially Thomsen (1924) and Koch (1980) (2008).

[6] See Basso (2018: 33-47) and Winkel Holm (1995), (1998).

[7] *Hvorvidt kan den i vor Literatur i sin Tid førte Strid om den frie menneskelige Villies Realitet siges at have ført til et blivende og udtømmende videnskabeligt Resultat?* Cf. Thomsen (1924: 5).

Kant: now it is important to note that what Howitz criticized in the Danish law concerning the issue of the mental illness and ascribing responsibility was its Kantian assumption.

From a legal point of view, in the specific case of Denmark, it is worth to noting that already the *Danske Lov* [Danish Law] of King Christian V in 1683 (the first unitary Code of the Kingdom, which was the basic point of reference until the Danish penal reform of 1866) declared in the article 6-6-17 that a crime committed in state of *delirium* or *furor* could not be punishable, and even the jurists of the end of the 18th century agreed with this: "Nobody can or ought to be punished for a crime, unless at that time he had been *morally free*", and "nobody can be punished if his actions were not *free*"[8] (Nørregaard 1784a: 9, §§ 1008 and 1009).

The point here is to establish what precisely it means to be *morally free* and, especially, why a mad person is "not-free"? The most difficult issue in this sense seems to be to define the mad person as not being responsible for his own actions because he was "not free" to act according to a conscious will following a *rational* and *moral* criterion. But at this point the anthropological question arises: this question lies outside the medical field and essentially concerns the philosophical sphere. Thus, the philosophical sphere is the battlefield on which the Howitz-dispute was fought. And, in fact, this philosophical perspective defines the real originality of Howitz's treatise that, otherwise, would be just one among a number of similar writings on the issue imputability at the time. Among the other authors, worthy of mention are, for instance, the so-called French "alienists", like *Étienne-Jean* Georget (1795-1828) *in primis*, who between 1825 and 1826 wrote three treatises on this topic;[9] the German, Johann Christian August Heinroth (1773-1843), who as early as 1818 had claimed the necessity of a *psychisch gerichtlichen Medizin*[10] (judicial psychic medicine) and between 1825 and 1833 devoted two works to it;[11] and Adolph Henke (1775-1843), whose *Abhandlungen aus dem Gebiete der gerichtlichen Medicin* (Treatises on Judicial Medicine) of 1815 Howitz quotes in detail.

Therefore, Howitz's treatise has the merit of originality not only because, from a chronological point of view, it comes before many of the most im-

[8] "Ingen kan eller bor straffes for en Misgierning, med mindre han haver havt moralsk Frihed paa den Tid".

[9] Georget (1825), (1826a), (1826b). Beside Georget, see also Mathieu Orfila (1823) and especially François-Emmanuel Fodéré (1813), whom Howitz mentions in his treatise.

[10] Heinroth (1818: 250).

[11] Heinroth (1825; 1833).

portant writings on the theme of madness and imputability (and its author is thoroughly acquainted with the international scientific literature on mental illnesses), but also because it shows how at the root of the Danish clinical and legal reasoning of the time, there was the Kantian moral doctrine as it was presented in the *Metaphysische Anfangsgründe der Rechtslehre* (Metaphysical Elements of Justice), that is, the first part of *Die Metaphysik der Sitten*, (1797, Metaphysics of Morals). There the presentation of the concepts of *Wille* and *Willkür* (translated here as "Will" and "will", following the standard English translation) answered the need of conceiving the fundamental principles and concepts of morals *in concreto*.

Kant's system of morals is based on an idea of Will considered as "the faculty of desire whose internal ground of determination…is found in the reason of the subject" (AA VI: 213; Engl. tr: 12-13).[12] Accordingly, the Will was explained in it as the faculty of desire regarded not, as is will, in relation to action, but "rather in relation to the ground determining will to action. The Will itself has no determining ground; but insofar as it can determine will, it is *practical reason* itself". The will that can be determined by the *pure reason* – Kant concluded – is called free will, while the will that is only determined by inclination (sensible impulse, *stimulus*) would be animal will (*arbitrium brutum*). Human will, by contrast, is the kind of will that is *affected* but not *determined* by impulses; accordingly, "a part from an acquired facility with reason", it is not pure in itself, but it can nevertheless be determined to actions by pure Will. Kant writes: "*Freedom* of will is just the independence of determination by sensible impulses; this is the negative concept of freedom. The positive concept of freedom is that of the capacity of pure reason to be of itself practical" (AA VI: 213: Eng. tr.: 12-14).

[12] Howitz specifies that he refers exclusively to the *Metaphysics of Morals* because "it is, as far as I understand, the last of Kant's writings dedicated to this issue and therefore it should be considered as his definitive thought in this regard". For Kant's quotations, Howitz actually draws above all from explanatory works on the work of the German philosopher, such as Kiesewetter (1795), whose Danish translation of 1797 was widely used, cf. Koch (2003b): 46 and Schmid (17953). Anders Sandøe Ørsted (1824a: 127), among other things, reproaches Howitz for not having chosen the most suitable Kantian textual places to explain the doctrine of freedom: "Kant is not wrong, at the point on which Prof. Howitz especially dwells – *which, moreover, is by no means the best source to explain Kant's doctrine of freedom* – when he rejects the definition of freedom according to which the latter should consist in a faculty of choosing or not choosing the law, asserting that the possibility of deviating from the law is not a faculty, but a weakness". Between 1790 and 1800 Kantian philosophy had a decisive role in Denmark in academic life and in cultural life in general. However, discussions on Kant in the journals of the time show a predominant interest in moral philosophy. Ørsted and the other followers of Kantian philosophy were particularly concerned with refuting the doctrine of happiness, according to the dictates of the *Metaphysics of Morals*, cf. Thuborg (1951), Koch (2003: 34), and precisely on this point Ørsted returns in his objections to Howitz.

Freedom of will, anyway, cannot be defined – as it has been defined – as the ability to make a choice (*das Vermögen der Wahl*) for or against the law (*libertas indifferentiae*), Kant states, even though "will as *phenomenon* provides frequent examples of this experience" (AA VI: 226; Eng. tr.: 19). For we know freedom (as it first becomes manifest to us through the moral law) only as a *negative* property in us, namely, that of not being *necessitated* to act through any sensible determining grounds. But we cannot present freedom theoretically as a *noumenon*, that is, freedom regarded as the ability of the human being merely as an intelligence, and show how it can exercise constraint upon his sensible choice; freedom cannot therefore be presented as a positive property. But we can indeed see that, although experience shows that the human being as a sensible being is able to choose in opposition to as well as in conformity with the law, his freedom as an intelligible being cannot be defined by this, since appearances cannot make any supersensible object (such as free choice) understandable. We can also see that freedom can never be located in a rational subject's being able to choose in opposition to his (legislative) reason, even though experience proves often enough that this happens (though we still cannot comprehend how this is possible).[13]

According to Howitz (1824: 34), Kant is close to Pelagianism: the deduction of his conception of freedom from the existence of the moral law could be recognized in Celestius, a disciple of Pelagius, who stated: "If now man should be without sin, then he can be without sin, and if he cannot, then it could likewise not be obligatory". Howitz even accuses Kant of being "Ultra-Pelagian", since his view of freedom is not freedom to choose, a *libertas indifferentiae* between good and evil, but a "freedom of virtue", so to speak, a "*Dydsfrihed*", in Danish.[14] Therefore, a freedom that not only presupposes the possibility of the human being acting in a moral and rational way, but also the necessity of doing so, in accordance with the autonomy and the infallibility of the moral law.

Moving from such an idea of autonomy, that is, the idea of freedom as the ability of *pure* reason to be itself practical, the Danish jurisprudence of that time considered the foolish person to be "not-free". He was not-free because he was "unable to determine himself according to a *rational* basis". This was the definition of madness given by the above-mentioned German doctor and scientist A.C.H. Henke in his treatise of 1815. In the Chapter Six of his treatise,

[13] See AA VI: 226; Eng. tr.: 20: "*Nur das können wir wohl einsehen* […] *das die Freiheit nimmermehr darin gesetzt warden kann, daß das vernünftige Subject auch eine wider seine (gesetzgebende) Vernunft streitende Wahl treffen kann; wenn gleich die Erfahrung oft genug beweist, daß es geschieht (wovon wir doch die Möglichkeit nicht begreifen können*).

[14] The term is coined by Howitz and appears in the official Danish dictionary of Christian Molbech (1828-1833: 620), under the entry "Valg" (Choice).

Howitz (1824a: 26) considers this definition to be emblematic. The Dane's attention is directed towards the "rational basis" mentioned in this definition of madness because – especially according to the leading research of the French "alienists" of that time – it was now accepted that "there exists a kind of madness that does not have its basis in the *reason* disorder" (Howitz 1824a: 26). Pinel's so-called *folie raisonnante*[15] is an example of madness in which reason perfectly works, and in this direction, we could also mention the so-called "*moral* insanity" described by the English doctor James Cowles Prichard (1786-1848) in his *Treatise on Insanity and Other Disorders Affecting the Mind* of 1835. A foolish person, in this sense, would be perfectly able to conceal his madness, and, moreover, according to the definition of madness as "the inability to determine oneself according to a rational basis", we should also include among the "foolish" infants, drunken people or somnambulists. Then, if in the final analysis, to lack freedom means being unable to determine oneself according to a *rational* basis (and this expression in the Kantian system refers to the conclusions of the so-called practical *reason*, that is to the moral law), then the distinguishing trait of madness would be the lack of morality, something that is negated by experience, Howitz says, moreover, because there are a lot of cases of melancholiacs suffering from moral scruples.

Coming back to the point of departure: who are really the foolish? Are they "free"? But especially, what is freedom? Is the so-called "sane" human being free? In this context what Michel Foucault (1972: 533; Eng. tr.: 515) wrote in his *History of Madness* is meaningful: "The madness of the nineteenth century would tirelessly recount the winding journeys of *freedom*".

According to Howitz, the human being himself isn't free, since every human action is necessarily determined by a motivation that weighs more than another motivation, and the so-called rationality is nothing but a *capacitas motivorum* (the ability to weigh up, to evaluate motivations). Freedom as *capacitas motivorum*, Howitz says, should be freedom legally considered, a freedom that has nothing to do with morality. Following the English empiricists, Howitz talks about a *necessary* influence of the motivations: motivations, both the rational, good ones, those connected with public utility, and the bad ones, on the one hand, are the necessary product of an innate organization, and, on the other hand, depend on the circumstances so that "every choice, without exception, presupposes that one of the two moments is heavier than the other one. A determination of the will, independent from any motivations, would be like an effect without a cause, both in the sphere of the morals, and in the physical one" (1824a: II-III).

[15] See Georget (1825: 72).

Howitz's perspective is thus just the opposite of the Kantian one. As the doctor writes in the introduction of his treatise (1824a: II) : "My research on nature had indicated to me a direction that is incompatible with the Kantian doctrine on morals, thus instead of moving […] from the discrepancy between the physical world and the spiritual one […], drawing from one of the most secure bases of my thought, I decided to considered the human being in his actual unity, as an inhabitant of the planet Earth, subject to the properties of nature and completely like the other animal creatures, to which he is almost totally similar".

In this direction, according to Howitz (1824a: 2), madness would be a "limitation in the use of the faculty of reason, because of an affection of the material organs that cause the activity of the (so-called) mind". To be more precise, in Howitz's perspective it is not correct to talk about body and mind in a dualistic way, but rather it would be correct to express oneself in terms of material substratum and function. Then, since madness manifests itself by means of acts, "it consists in a lack of rational self-determination due to the physical cause". In this sense, the "rational self-determination" is nothing but the faculty of deliberately acting or not acting, determined by "rational representations". These representations are namely "rational" – they are ideas – in so far as they *correspond* to real objects, and they work as rational motivations when they make it possible to expect the advantageous consequences of an action for the acting subject or for anybody else. Here Howitz (1824a: 16-17) explains what he means with the conformity between "representation" and its "real object"; this conformity namely means "accordance":

> Such a conformity is attested by the accordance among human beings with regard to the sensible impressions. The human sane understanding abstracts by experience a certain rule in thought and in desire, and in such a representation are subsumed all the ends of society and the way in which human beings interact with one another. At the root of this rule there is the accord among human beings concerning the sensible impressions, the judgement on things and on what is attractive and repulsive. Probably this accord is not perfect, but if it didn't exist, it wouldn't be possible to recognize any basic equality above the different individual shades due to the different age, sex, temperament, innate aptitude, education, state, nation, historical epoch, etc. Without this basic equality there wouldn't be any common experience, any public opinion, truth, common good or evil, and human beings would have never realized anything together.
>
> If we run into a man who lacks such a basic equality – if, for instance, he hears voices that others do not hear, he sees something shining that the others see to be dark, he is afraid of consequences where according to all the others there are no causal relationship, or vice versa he does not recognize an evident danger or he is attracted by

something that those like him judge disgusting…, – if we run into a man like that, we will consider him wrong, and if he does not accept to be persuaded, we will consider him mad, *mente alienatus*, as the ancients said, *aliené* in French, that is, someone who is deviant, or led astray from a condition that, even if it is not primitive, is, however, considered the normal condition for the human being in general.

Of course, we cannot say that a man who lacks of rational representation is necessarily mad. It might be possible that his rational representations have a weak influence compared to his sensitivity, and in this case he would act in a thoughtless or immoral way. Nevertheless, the man acts in an insane way only in the case when the imperfections of his understanding or the lack of influence of the latter are due to a physical illness.

Howitz's moral perspective is actually a utilitarian view combined with a deterministic view of nature, where natural necessity, however, does not conflict with an idea of freedom as possibility to determine oneself rationally, but it is rather opposed to constraint.

What Howitz rejects is the philosophical view that considers human being as a *homo duplex*, made from the one side of a soul having will and from the other of weak flesh, a sensitivity that can be yet defeated by a strong will. This is the idea that was at the root of the liberal-conservative view of the law.[16]

"When Galileo presented his grand discovery" – Howitz comments in another essay dated 1825 (1825: 51), that is, in the last act of the dispute – "he found himself forced to fight the general prejudice according to which the celestial bodies belonged to another order of being in comparison with the known world, and they weren't subject to the earthly law".

The natural condition of the human being considered in his entirety – so also in his materiality – that is, the state of the human being as a natural being, the contingent configuration of his components (as they would arrange themselves, for instance in the case of an illness) can have an invincible effect on his will. Since this effect manifests itself with a *natural necessity* that excludes imputability, this necessity cannot be understood as we usually understand sensitivity from a philosophical, moral, juridical or religious point of view, that is, as something that can be overcome by free will. Foucault (1972: 533; Eng. tr.: 514) would have commented a century and half later: "The madman was freed from his association with crime and evil only to be locked into the rigorous mechanisms of a determinism. He was only completely innocent in the absolute of a non-freedom".

Howitz (1824a: 64) opposes to the view of a *homo duplex*, the idea of a *homo*

[16] See Holst Scherg (2006: 84-86).

triplex, whose will can deviate from the optimum conditions not only because of irrationality or immorality, but because of something physical, a brain illness: this affection is called madness. "The defect of the organization cannot be removed by will, it cannot be removed by motivations", and especially "it cannot be changed by the series of causes that we call spiritual".

Nevertheless, Howitz insists that it is important not to confuse the doctrine of determinism with fatalism. This happened on occasion of the publication of Étienne-Jean Georget's writing on madness and imputability of 1826, *Discussion médico-légale sur la folie ou aliénation mentale*. Since he refused the conviction of the murderer if he was insane – that is, if he was affected by a "lésion de jugement" or a "lésion morale (Georget 1825: 10, 15), the French physician was accused by an anonymous writer in the *Journal des débats* (18th February 1826) of encouraging fatalism because he would have defended an equivalence between murder and illness that is totally incompatible with the moral laws. This objection had also been raised against Howitz, even if from a more pragmatic point of view. Specifically, Anders Sandøe Ørsted, in a work that followed Howitz's treatise, *Et Par Ord i Anledning af den foranstaaende Afhandling* (A Few Words on the Occasion of the Aforementioned Treatise), objected that the meaning of the sentence as a deterrent, would have risked being lost if the popular belief had turned to fatalist positions (Ørsted 1824a: 135).[17] Moreover, if a crime could have been judged as something inevitable under particular physical conditions, one would have considered vain every effort towards morality. This was actually already an idea at its planning stage if we consider, for instance, the words of the Belgian Adolphe Quetelet (1796-1874): "We are able to enumerate in advance how many individuals will stain their hands with the blood of their fellow creatures, how many will be forgers, how many poisoners, pretty nearly as one can enumerate in advance the numbers of births and deaths which will take place" (Quetelet 1831: 80-81).[18] It is in these years indeed that we see the beginnings of *anatomical pathology and the use of statistics in medicine*.

The fear of Howitz was thus the possibility of considering *every* crime as justifiable due to a lack of freedom because of a physical disease. Nothing could be more wrong.

In order to examine the question, the presence of a doctor was necessary who could recognize the connection between a so-called "moral lesion" and a brain affection, according also with the medical policy of the time.[19]

[17] Cf. also Waaben (1997: 45-46).

[18] Cf. Radzinowicz (1966: 36).

[19] Cf. Georget (1821: 139-140): "Le cerveau est le siége immédiat, la cause organique essentielle,

This was what Howitz (1824a: 15) claimed in the fourth chapter of his work. He claimed that the doctors should always be consulted in uncertain cases, and their final judgement concerning mental insanity should never be invalidated by any court of law. But at that time the jurisprudence did not have any rule about the necessity of consulting a doctor, and in such cases the decision was simply delegated to the judge (Ørsted 1815-1822: 123). As Michel Foucault states in his lectures at the Collège de France in 1973-1974, "we see a very strange process in the courts in which doctors – who were not called on by the prosecutor or by the president of the court, and often not even by the lawyers – gave their opinion on a crime and, as it were, tried to claim the crime for mental illness itself" (Foucault 2003: 249; Eng. tr.: 249).

Foucault explains this process with political reasons rather than scientific or juridical, and this is interesting for us, because it will be an accusation that also was levelled against Howitz by his contemporaries, since the doctor's claim to have the final word was considered a claim to power. This is another thorny subject raised by Howitz.

The objection that Ørsted made against Howitz in this case was the impossibility in many cases of tracing the presence of a brain affection before the autopsy and sometimes even after it. In this last case one could say that there would actually be a "leap" between the empirical observation of a deviant behavior and the conclusion that its origin lies in a physical disease, a kind of belief. Nevertheless, Howitz was totally sure about the deterministic theory, even if he was not a follower of phrenology, which also in Denmark had many followers, among them the physician Carl Otto (1795-1879), who was actually the only one to side with Howitz when the treatise of 1824 was published. In the wider debate on the subject there were even some physicians – such as Joachim Dietrich Brandis (1762-1845) – who were critical of their own group. Again in the *Juridisk Tidsskrift* in 1824, Brandis stated in his essay *Om den Juridiske Bestemmelse af Afsindighed* (On the Juridical Determination of Madness) that the role of the doctors in a trial in which a mad person is involved can be only as a consultant, while the final decision should be taken only on the basis of a positive law (Brandis 1824: 206). But Howitz, for his part, is firm in his position. Although no follower of phrenology, he was an attentive observer of the most recent progress of morbid anatomy. The third chapter of Howitz's essay tries to demonstrate that madness is *always* determined by a physical illness, and he shows that this thesis is supported by the research of scientists such as

l'instrument principal de tous les phénomènes intellectuels, des sensations, des combinaisons de l'esprit, des passions, des affections, le point de départ des mouvements volontaires, enfin de toutes les opérations de l'organisme qui se font avec conscience".

Spurzheim (1776-1832) – the former pupil and collaborator of Gall – Nasse (1778-1851), Vering (1773-1829) and, of course, Georget.

The publication of Howitz's treatise followed the occurrence of some real episodes of doubtful cases of madness, in which it seemed there were discrepancies between madness as a psychic condition and its somatic causes. One particular case was so odd that it had been presented at the Royal Danish Society some years earlier by the physician Johan Daniel Herholdt (1764-1836).[20] Herholdt's case dealt with the strange sickness of a 26-year-old patient, Rachel Hertz, who showed various odd symptoms such as temporary paralysis, insensibility, insomnia, urine retention etc. without any evident physical affliction. In some periods, the symptoms mysteriously disappeared, but only for a while. Dr. Herholdt and his assistant were unable to diagnose the disease. Some years later Rachel began suffering pains in her abdomen. Herholdt examined her and found a hard lump causing unendurable pain at the touch of it. When he cut into the tumor he discovered inside a black sewing needle. In the following years Dr. Herholdt cut out no less than 389 needles from her abdomen. Rachel Hertz became a sensational story for the newspapers. People in Copenhagen talked about "the sewing needle girl", and she became known even abroad.[21] The patient had explained that many years previous she had swallowed a needle case. She was a very clever girl and was perfectly able to reason and discuss with other people. Since, due to other reasons that we cannot go into here, the case was inexplicable, the doctors decided to observe the girl secretly, and they discovered that she was a pretender, as they said, who simulates her symptoms and that she had put the needles under skin by herself. Some years later Breuer and Freud would have easily defined this as a typical case of hysteria. The patience of course was not a liar, even if it was not possible to find the organic causes of her disease.

Dr. Herholdt expressed his wonder in 1826 in his private journal asking: "Does a kind of madness that does not have its basis in the perturbation of the reason really exist?" These words are a quotation from Chapter Five of

[20] *Udtog af Dagbøger over Rachel Hertz's Sygdomme i Aarene 1807-1826: med tilføiede Bemaerkninger* [Excerpt of Diaries of Rachel Hertz's Diseases in the Years 1807-1826: with Remarks Added], Gyldendalske Boghandling, Copenhagen 1826, was translated also in German in the same year, under the title *Auszüge aus den über die Krankheiten der (Copenhagener Nadelschluckerin) Rachel Herz, während der Jahre 1807-1826 geführten Tagebüchern*, Seidelin, Copenhagen. Four years earlier Herholdt had documented the course of the young girl's illness in his work *Observatio de affectibus morbosis virginis Havniensis cui plurimæ acus e variis corporis partibus excisæ et extractæ sunt*, Havniæ, 1826. The most complete monograph on this topic is Michelsen's (1989). See in particular pages 20-36. See also Christensen (2009: 727-741, Eng. tr. 2013: 415-423).

[21] See Christensen 2009: 728-29; Eng. tr.: 416.

Howitz's treatise (1824a: 23). Where lies the *discrimen* between madness and mental health if it is not in the correct use of the faculty of reason? What is the boundary between the two states? And especially, who is the legitimate authority to establish it if in the end the only possible *discrimen* is the decision? This is another question that Howitz puts to the jurists and thus to the philosophers when they consider freedom as an ontological (transcendental) condition, which does not allow of gradations, since it is or not, *aut Caesar out nihil*. Since, according to Howitz, freedom is simply a *capacitas motivorum*, there are different degrees in it, a gradation that depends on the condition of the whole that the man is.

The accusation that Howitz thus leveled against the (Kantian) jurists was the claim to recognize *a priori* "a clear line between madness and sanity" (1824a: I) without acknowledging the existence of a series of problematic intermediate stages. We find here in a sense the Greek paradox of the sorites that the Skeptics used in order to demonstrate the impossibility of distinguishing truth from falsity: at what point, when moving one grain of sand after another, can I say that there is no longer some grains of sand but a mound? Is the mound something that is ontologically different from a grain? If it is not, the point is that one has to "invent" the reality of the concept of mound.

If such a clear line does not exist – Howitz says – is it fair that the law has to establish it? Moving from experience, Howitz states, we see that there exist degrees in freedom as *capacitas motivorum*. Moreover, the physician adds, the extremely wide classification of mental disorders is due to the impossibility of constraining the actuality into the boundary line established by the methodical intellect. The normal condition of the brain, as well as that of the body in general is never absolutely perfect, but always shades off into disease (Howitz 1824a: 63).

In distinguishing between the imputability and non-imputability, freedom and non-freedom, possibility and impossibility of a psychological constraint, madness and sanity of human understanding – Howitz states – we arrive at concepts that can be useful when we refer to the distant extremes of the opposite state. But the situation is different if we consider the many degrees and we compare the minor madman to the so-called "free man", "whose intellect is strongly restricted by a congenital organization, blind with prejudices, clouded with drunkenness or depression" (1824a: 79-80).

The case of Rachel Hertz was well known in Copenhagen and raised a great debate, but it was another clinical case that directly led to the decision by the young professor of forensic medicine to write his essay in 1824. This case is certainly less striking than that of the "sewing needle girl", but perhaps more problematic since now what is at issue is a case of attempted murder.

The young maid Anne Marie Lorentzen was eighteen years old in November 1821, when she tried to strangle the old Mrs. Bagger, whom she worked for in Copenhagen, in her sleep with a handkerchief. The murder failed because after hearing the commotion, the landlord's daughter downstairs decided to go into the Bagger house, thus interrupting Lorentzen's action. Anne Marie then, gripped by anguish and, as if awakened by the state of exaltation in which she found herself, fled to Kastellet to commit suicide, but instead of making the extreme gesture, she finally resolved to go to the police. She was then taken to the General Hospital, where she was treated for a week and then arrested. The young woman did not remember what had happened at the bed of Mrs. Bagger, who had been found by the landlady on the ground. Lorentzen confined herself to declaring that she felt that she had been the prey of an "invincible inner strength". The case was submitted to the College of Health in Copenhagen, where the medical record of the young woman was studied and her medical history reconstructed.[22] The College – Howitz in the first place – expressed itself by declaring that Lorentzen had acted in the instant of the attempted murder in a state of "unconscious fury" attributable to "a physical illness" and therefore was not attributable. On May 13th, 1822, the city court acquitted the young woman, but with a reservation: Her present state (at the time of her absolution) was considered "healthy". The altered state that had caused her to commit the criminal act had been judged to have passed, and therefore legally it would not have been necessary to apply article 1-19-7 of the *Danske Lov* which dictated that mentally ill people dangerous for the community would have to remain in police custody. Nevertheless, the girl was placed under the supervision of the authorities. The Supreme Court ratified the sentence on June 20, 1822, but found itself forced to contact the Danish Chancellery immediately afterwards because it had not been able to define in what exactly the "surveillance" should have consisted. They then returned to consult the College of Health, which did nothing but request a solemn declaration from Lorentzen which assured that she would immediately turn to a doctor in contact with the police if she perceived anew the signs of the illness that preceded her criminal action. When it came to voting in the College, in July 1822, Howitz wrote in his own judgment that Lorentzen should have been "treated by the authority with all possible respect and humanity", so that she could convince herself of the opportunity to go to the doctor and to the police "without fear, like her guardians and her defenders". It was the first time in the history of Danish justice that the authorities expressed themselves in such compassionate and solicitous terms about a criminal (Waaben 1997: 38).

[22] See the recordings of the trial in Lange (1822).

At the beginning of his 1824 treatise, in fact, Howitz listed the motivations that had led him to try his hand at a theoretical treatment of the problem, which had to be examined in its ideological components in order to root out the prejudices that founded a certain way of proceeding in criminal proceedings. Following the discussion of the Lorentzen case examined at the College of Health, in fact, Howitz declared that he had had the opportunity to hear various opinions on the subject both from the doctors and from the jurists. He noticed that the latter placed a too strict boundary around the madness, ignoring the intermediate degrees, the gradual transition between madness and "normal human reasoning". However, not satisfied with having empirically demonstrated his position and shown the Kantian terrain at the base of the jurists' prejudices, he considered it appropriate to enter into a "metaphysical and moral investigation", that was the beginning of the subsequent dispute. Another essential point for Howitz was to defend himself against the accusation of being a follower the so-called theory of pleasure, since for him the pursuit of pleasure, or of what is beneficial to oneself, did not coincide with selfishness. Kant's objections to the eudemonistic doctrine are known, in particular that of Wolff or Helvétius, since this doctrine would deny the autonomy of morality by making it heteronomous. For Howitz, however, who expressly mentions Bentham –"one of the greatest living jurists of England" (Howitz 1824a: 60 footnote) – and his *Traités de législation civile & pénale* of 1802, morality is the doctrine of happiness for social life, and virtue is the art of living happily among men. Therefore, the "morally good" is what springs from the reasons of the common good. The correspondence between will and instinct of happiness should therefore never be denied, unless the instinct of happiness is exchanged with low selfishness and sensuality, "since it has been forgotten that in man there are also instincts like the sense of humanity, motherhood, honor, whose satisfaction requires the sacrifice of sensitive desires. The instinct of happiness together with reason must consider the most remote consequences; it must be determined by hope and fear, and one has to remember the old rule *nocet empta dolore voluptas*" [pleasure bought by pain is injurious][23] Howitz 1824a: 55).

In short, Howitz declared – now quoting Locke: "Where is that practical truth that is universally received, without doubt or question, as it must be if innate?" (Locke 1690: I, II) – that it seems indubitable that morality is founded on the coexistence of men and their mutual social relationships, while the categorical imperatives would be absolutely mute if man were still living in *statu solitudinis*. This would not conflict with Howitz's deterministic view of free-

[23] Horatio, *Epistulae*, I, 2, 55.

dom understood as *capacitas motivorum*, where man's love for good or for what is of public utility (moral law) as well as his propensity to strive in the direction of eternal life would be nothing but a particular kind of motive alongside sensitive ones. It is evident that in this case the physician, far from proposing a deontological order and trying to found it, merely emphasizes the functional structure of his necessitarian theory of motivations.

It is especially in the chapters §§ 5, 11 and 12 of his treatise that Howitz discusses the problem of the effective identification of a mental pathology and related nosology. It is difficult for an action carried out under the aegis of madness to be a rational action, he writes, but this does not mean that every irrational action is to be blamed for madness; otherwise, even simple-mindedness and strong passion would both be faces of madness. The determining factor for Howitz always remains in the cause of criminal action, or *physical* illness, even if this not always (immediately) provable. Although some types of real madness due to physical causes depend on a complete lack of reason – idiotism – or on a fixation on certain ideas, there are other kinds that seem to depend more on the *lack of influence* of reason than on its defective nature. In that case reason would suggest the right thing, but its advice would either be too weak or would intentionally not be followed. This is the case, for example, of that pathological "gloominess" that manifests itself without presenting false ideas, but only with transformed feelings, tendencies and passions. These so-called fixed *passions* (note the difference from fixed *ideas*) would also be caused, for Howitz, by physical defects in the brain and can even be more dangerous than fixed ideas. For example, the so-called *iracundia morbosa* belongs to this group of afflictions, a condition in which a person would be driven to an extraordinarily intense state of anger and consequently to violence for the most insignificant reasons. Another type of affection that falls within the group is the *sine mania sine delirio* or the *furor transitorius* described by Pinel (1809), which consists of a similar access of temporary anger accompanied by the desire to destroy and kill. People in this state are often not aware and sometimes even ask those close to them to move away: in other words, there is no other mental disorder except for the urgent homicidal desire. According to Pinel, it is still be quite simple to be able to distinguish these outbursts of fury from true moral depravity because in general they manifest contradictoriness with respect to the habitual character of the person involved. In the case of *iracundia morbosa*, by contrast, this contradiction does not manifest itself; instead there is an agreement between the innate temperament and the insurgent illness. And here Howitz proposes to distinguish between *constitutional*, idiopathic, and *accidental* madness, a

distinction that does not seem to find any comparison in any of the authors he read: in short, *iracundia morbosa* would be a kind of "constitutional" madness, while *mania sine delirio* would be an "accidental" illness. Moreover, the more "physical" the instinct is whose satisfaction is coveted by madness, the more evident it is that the cause of the illness should be recognized not in a lack of reason, but in its non-use as a motivation for action: that is, instinct influences the will and decides, despite the objections of reason. And this would be the case with diseases such as nymphomania, satyriasis, *appetitus gravidarum* and dipsomania (alcoholism understood as physical illness). In the case of similar states, Howitz states that he thought it more appropriate to place the essence of madness in the lack of rational self-determination due to physical illness, than in the lack of reason *tout court*. The latter case, which has already been seen in part, is that of the *folie raisonnante*.

In short, we see a proliferation of classifications due precisely to the impossibility – and Howitz is perfectly aware of this – of bringing reality into the "signs of demarcation established by the ordering intellect" (Howitz 1824: 79). The normal state of health of the brain, as well as of the whole body, is never absolutely perfect, but "always fades in disease" (Howitz 1824: 65). When one distinguishes between imputability and non-imputability, freedom and non-freedom, the possibility and impossibility of a psychological constraint, madness and normality of the human mind, such concepts, declares Howitz, may very well have a reality and be perfectly usable when referring to "the far extremes of opposing states"; however, the situation is quite different if we observe the intervening degrees and confront the slightly insane with the so-called "free man" "whose intellect is strongly limited by a congenital organization, blinded by prejudices, clouded by drunkenness, pathologically demoralized" (Howitz 1824: 80). It is inevitable to encounter difficulties in wanting to establish borders scientifically, the same difficulties encountered by the "systematic philosopher of nature in the passage from plants to animals or from alkalis to soils" (ibid.). It is on this point that, as we have seen, Howitz rages against the jurists, insofar as they would not consider these gradual degrees. In other words, there are mental states which, due to limited freedom (i.e., self-determination according to the normal human intellectual capacity), could be considered analogous to madness. And since the latter is not a reason for imputability, then one should believe that the states related to it should appear to be subject to less imputability. In reality, if in the past they were considered as extenuating reasons of punishment, modern criminalists, for Howitz, would now deny them any influence in attenuating the positive penal code: "According to them, degrees in the freedom of the act are not a scale of measurement of punishment, as are degrees of immorality". Here Howitz quotes

the *Eunomy* by AS Ørsted, the *Lehrbuch des gemeinen in Deutschland gültigen peinlichen Rechts* by P.J.A. Feuerbach of 1801, as well as the Danish decree of 18 December 1767, which already contained the statement that "no middle way can be assumed between this lack of use of reason, which completely excludes the application of the penalty and status of the person in whom such a penalty must be applied" (Ørsted 1804-1812: § 77).

In this way a "clear difference" would be established between all those who, based on the words of the decree, "are defrauded of the use of the intellect" and those who are in possession of it, or between "absolute impunity, on the one hand, and the most severe imputability, on the other. The border is called *capacitas motivorum*, which is supposed to be now present, and now completely missing", in the same way that, according to the Kantian doctrine, "freedom is either completely present or not at all" (Howitz 1824: 46). The doctor knows from experience that in many points, far from being clearly distinct conditions, they become confused, and "there are degrees from a man in an asylum to a man in a sociable condition. And vice versa there are degrees between the mental state that allows the possession of civil rights and that which denies them, and thus it is always impossible to determine whether legal freedom (*capacitas motivorum*) was present at the moment a crime was committed. At the same time there is a great similarity between the many for the whom the threat of punishment is put in place and the many others who avoided this penalty due to the absence of *capacitas motivorum*" (Howitz 1824: 92).

Then Howitz proceeds to closely observe some of these intermediate steps or stages between madness and normal human reasoning. He lists six categories: 1) cases of intermittent madness", or so-called *lucida intervalla*; 2) the *follia partialis*, which is characterized by fixed ideas; 3) the lowest degrees of each type of madness; 4) the stages of development of the disease and the recovery from the disease itself (see the case of Anne-Marie Lorentzen); 5) *status semisopidus*, or half-sleep, hypnagogic states and other similar transient delusions of physical origin; 6) vices and passions of the so-called "free" man.

As a premise of his phenomenology, Howitz highlights how the common prejudice in those who have no empirical or theoretical familiarity with madness is the belief that it is always "total", that is to say, always and in any case as *dementia* or *manias generalis et absoluta*. In other words, it is believed that a person who once performed acts that authorize us to call him mentally ill, will persist in a state of unbroken madness until death. However, it is enough to visit an asylum for only an hour, Howitz warns, to realize that even those who have been declared mentally ill are not always and in all respects insane. There are periodic states when the illness stops in an unpredictable way (intermittent), and during such a state the two poles of the interval ei-

ther do not express themselves at all (*intervalla lucida* or *dilucida*) or do so to very low degree (*intervalla obscura*). Thus, these people cannot be reasonably exempted from the imputability for all their deeds: this is why it is necessary for a doctor who knows the disease to evaluate it. In short, one imagines the lucid intervals as "isolated bright spells in a long night", but in reality the situation is exactly the opposite, that is, there are numerous dangerously mentally ill people, declared as such only following a paroxysmal attack lasting some days or a few weeks, but then for months or a whole year they are calm and harmless and sometimes even in full possession of their faculties. Of course, Howitz continues, it cannot be said that the judges did not take into consideration periodic madness and lucid intervals; indeed they seem to have placed their attention "exclusively" on this sort of interval, is if it were a bridge capable of leading by madness to freedom and normal human reasoning. This can be seen by looking at the Danish law and its interpreters: "Those who are endowed with the use of the intellect at intervals, in those moments should be considered like other rational men, or as beings able to act freely; therefore, they should be punished for their misdeeds, while in other periods they should be placed in the class of fools" (see Nørregaard 1784a: 89 and Hurtigkarl: 1813-1820: I, 8).

Another frequent mistake concerning the idea of periodic madness, Howitz continues to warn, is to consider every *lucidum intervallum* as absolute, while often such *lucida intervalla* consist in an attenuated degree of disease. In short, the *insania remittens* – an attenuated degree of disease – would be exchanged with the *insania intermittens*, or a temporary cessation of the disease. At the level of jurisprudence, to put these semi-lucid intervals in the class of those where the perfect normal state takes place, would be very unfair, says Howitz. However, modern criminalists have denied the *obscura intervalla* in a legal sense, and when such an event instead is verified in a crime, they considered it necessary to have to "do violence to the truth", assuming that either there was no interval or that there was an *intervallum lucidum*. But, Howitz says, let's not forget that the life of the defendant depends on such a decision.

The second category of psychological states in which the infirm of mind could be considered *capax motivorum* is that of *insania partialis*, the so-called *monomania*. Most of the mentally ill who are subject to a dominant fixed idea are able to reason and judge correctly everything that has nothing to do with their fixed idea.

A third objection to the assertion that the mentally ill must be considered unconditionally as a *non capaces motivorum* concerns the slightest degrees of madness, which are encountered both inside and outside civil society, in mental hospitals. This applies in the first place to that class of patients who call

themselves "infantiles" or "imbeciles" and whose state is considered as a more or less generic paralysis of mental faculties. Whether it is a congenital defect, which has taken over with old age, which is the consequence of some pathology or due to an accident, if it occurs in mild degrees, the subjects who are affected show signs of being able to distinguish what is lawful from what is not and of being motivated by fear of punishment. This is also true for another class of semi-infirm minds: the depressed and the melancholic, who, generally, are in asylums because of a so-called *raptus melancholicus*, but in such cases the disease is attenuated.

So what is now the "psychological criterion" on the basis of which we can distinguish these subjects from the depressed and the melancholy who are instead outside the asylums and who – on basis of the Decree of 18 December 1767 – were judged to possess the use of reason? In reality they are both equally vulnernable to the possibility of *raptus melancholicus* and to committing or attempting to commit crimes. To say that both should be now considered in a state of *lucidum intervallum* is not correct, because the fixed idea in them remains, and anger is not only a degeneration of the existing pathological mental state. To define this state as an *intervallum obscurum* would seem closer to the truth, but not in the view of modern criminalists, for whom this expression should be abandoned and its concept reduced to non-freedom or imputable freedom, an extremely arduous operation for the honest judge.

According to Howitz, the melancholic is not the only partially insane person that in the milder degrees of the disease could be considered *capaces motivorum*, but this applies to almost all types of madness and above all to partial madness. The fourth case, which we have already seen with Miss Lorentzen, shows, on the one hand, the development of the disease and, on the other, the convalescence from it. Madness rarely breaks out without warning, but only the "expert of the human soul" is able to listen to these warnings and prevent their consequences.

In the fifth case we speak of sleep as a condition of non-freedom, and consequently the moment in which the individual wakes up or falls asleep is a state of transition from freedom to non-freedom. The *status semisopidus* therefore entails a state of delirium. It is rare for this condition to become an object of reflection for the judge, but there are examples of assassinations carried out in a hypnagogic state, or where "terrifying images coming from the realm of dreams have been projected into people who actually surrounded the subject in question. Then, these obscure sensitive impressions began to recall the subject's conscience from its lethargy" (Howitz 1824: 108). This state, Howitz points out, is not comparable to that of the sleepwalker or to magnetic sleep since these two cases are conditions of total non-freedom, while the first case is

a transition: it is more similar to the incipient or vanishing intoxication caused by alcohol and narcotic substances, or to the state that immediately follows convulsions or coma. Suicide is often carried out in this state.

The sixth, and last, case is presented by Howitz in an almost provocative way, to convince his readers how incorrect it is to establish a clear boundary between madness and non-madness. The same so-called "normality" is not in fact exempt from paroxysmal excesses in which the subject, in the moment of action, is probably not *capax motivorum* at all. We refer to jealousy, for example, or the fury of revenge. "I do not believe", Howitz concludes, "that the threat may have some influence when the passion has reached such intensity" (ibid.: 112). And in these cases of excess Howitz also includes the love that charms reason or the religious enthusiasm that drives us to do something against reason, as in the case of the African sorcerers who throw themselves from the heights with the intention of injuring themselves. Finally, Howitz closes in an almost moralizing way: "Among men in society, we come across fixed ideas at least as often as we encounter in the dominant affects, and they also have the same analogy with *insania partialis*. What are our factiousness, our obsessions, our innumerable prejudices and the false conclusions that decide our way of acting despite all the objections, if not fixed ideas? What are our infinite kinds of superstitions, if not the same thing?" (*ibid*.: 116).

The "*Howitzfejde*" would cause a great deal of controversy even many years after the untimely death of its initiator in April 1826. It is an emblematic episode, and, therefore, it is worthy of attention: through it, it is possible to perceive a *Zeitgeist* that is embodied in the institutional reality of an age and a country, informing the choices of a society and consequently its material structure, to the point of rendering the physical organism of this society indistinguishable and the principles that inform it. If therefore this "*Zeitorganismus*", as one can rightly define it, develops in its inseparable components following a global movement, it would be the exclusive task of the philosopher to analyze it, in an attempt to reach an awareness about the way in which everything interacts and gives life to those changes that define a world as properly human.

Ingrid Basso
ingrid.basso@unicatt.it
Università Cattolica del Sacro Cuore, Milano

References

The Penal Code of France translated into English with a Preliminary Dissertation and Notes, 1819, Butterworth, Baldwin, Cradok and Joy, London.

Aagaard Olesen, Tonny, 2003, *Kierkegaards Schelling. Eine historische Einführung*, in Jochem Henningfeld and Jon Stewart (eds.), *Kierkegaard und Schelling. Freiheit, Angst und Wirklichkeit*, De Gruyter, Berlin-New York: 19-21.

Algreen-Ussing, Tage, 1831, *Haandbog i den danske Kriminalret*, Kjöbenhavn, Hostrup Schultz.

Basso, Ingrid, 2007, *Kierkegaard uditore di Schelling. Tracce della filosofia schellinghiana nell'opera di Søren Kierkegaard*, Mimesis, Milano.

Basso, Ingrid, 2017, "Su follia e libertà del volere. La 'disputa Howitz' nella Danimarca del primo Ottocento", Saggio introduttivo a Frantz Gotthard Howitz, *Su follia e imputabilità. Un contributo alla psicologia e al diritto*, a cura di I. Basso, Mimesis, Milano: 16-32

Basso, Ingrid, 2018, "La 'forma negativa della libertà infinita'. Frantz Gotthard Howitz negli scritti di Søren Kierkegaard", in Amoroso L., Garaventa R., Rocca E. (a cura di), *Kierkegaard e il suo tempo*, in *NotaBene. Quaderni di Studi Kierkegaardiani*, 10: 33-47.

Bentham, Jeremy, 1802, *Traités de législation civile & pénale*, Bessange, Masson et Besson, Paris.

Brandis, Joachim Dietrich, 1824, "Om den juridiske Bestemmelse af Afsindighed", in *Juridisk Tidsskrift*, 8, 2: 205-221.

Christensen, Dan Charly, 2009, *Naturens tankelæser: en biografi om Hans Christian Ørsted*, Museum Tusculanum, Københavns Universitet, København, bd. II, Kap. 2: *Brødrene Ørsted i Howitz-fejden*: 727-741; Eng. tr. *Hans Christian Ørsted, Reading Nature's Mind*, Oxford University Press, Oxford 2013: 415-423.

Fodéré, François-Emmanuel, 1813, *Traité de medicine légale et d'hygiène publique ou de police de santé*, Mame, Paris.

Foucault, Michel, 1969, *L'Archéologie du savoir*, Gallimard, Paris; Eng. tr. by Alan M. Sheridan Smith 1976, *The Archaelogy of Knowledge*, Harper & Row, New York.

Foucault, Michel, 1972[2] (1961), *Histoire de la folie à l'âge classique*, Gallimard, Paris; Eng. tr. by Jonathan Murphy and Jean Khalfa, *History of Madness*, Routledge, London-NewYork 2006.

Foucault, Michel, 2003, *Le pouvoir psychiatrique. Cours au Collège de France (1973-1974)*, éd. établie sous la dir. de F. Ewald et A. Fontana par J. Lagrange, Gallimard-Seuil, Paris 2003, Eng. tr. by Graham I. Davidson, *Psychiatric Power. Lectures at the Collège de France*, edited by Jacques Lagrange, general editors: Francois Ewald and Alessandro Fontana, English Series Editor: Arnold I. Davidson, Palgrave MacMillan, New York 2006.

Georget, Étienne-Jean, 1820, *De la folie. Considérations sur cette maladie. Son siége et ses symptômes; la nature et le mod d'action de ses causes; sa march et ses terminaisons; les différences qui la distinguent du délire aigu; les moyens de traitement qui lui con-*

viennent; suivies de recherches cadavériques, Crevot, Paris.

Georget, Étienne-Jean, 1825, *Examen médical des procès criminels des nommés Léger, Feldtmann, Lecouffe, Jean-Pierre et Papavoine, dans lesquels l'aliénation mentale a été alléguée comme moyen de défense, suivi de quelques considérations médico-légales sur la liberté morale*, Migneret, Paris.

Georget, Étienne-Jean, 1826a, *Discussion médico-légale sur la folie ou aliénation mentale, suivie de l'examen du procès criminel d'Henriette Cornier, et de plusieurs autres procès dans lesquels cette maladie a été alléguée comme moyen de défense*, Migneret, Paris.

Georget, Étienne-Jean, 1826b, *Nouvelle discussion médico-légale sur la folie ou aliénation mentale, suivie de l'examen de plusieurs procès criminels dans lesquels cette maladie a été alléguée comme moyen de défense*, Migneret, Paris.

Heinroth, Johan Christian August, 1818, *Lehrbuch der Störungen des Seelenslebens oder der Seelenstörungen und ihrer Behandlung*, Vogel, Leipzig, 2 voll.

Heinroth, Johan Christian August, 1825, *System der psychisch gerichtlichen Medizin*, Hartmann, Leipzig.

Heinroth, Johan Christian August, 1833, *Grundzüge der Criminal-Psychologie oder: die Theorie des Bösen in ihrer Anwendung auf die Criminal-Rechtpflege*, Dümmler, Berlin.

Henke, Adolph Christian Heinrich, 1815-1834, *Abhandlungen aus dem Gebiete der gerichtlichen Medicin, als Erläuterungen zu dem Lehrbuche der gerichtlichen Medicin*, Kunz, Bamberg.

Henriksen, Aage; Nielsen, Erik; Wentzel, Knud (eds.), 1975, "*Howitz-fejden*", in *Ideologihistorie*, Vol. 1: *Organismetænkningen i dansk litteratur 1770-1870*, Tabula/Fremad: 105.

Herholdt, Johan Daniel, 1822, *Observatio de affectibus morbosis virginis Havniensis cui plurimæ acus e variis corporis partibus excisæ et extractæ sunt*, Havniæ.

Herholdt, Johan Daniel, 1826, *Dagbøger over Rachel Hertz's Sygdomme i Aarene 1807-1826; med tilføiede Bemærkninger*, ed. J.D. Herholdt. København.

Holst Scherg, Rune, 2006, "Synd, forbrydelser og laster: forbryderen i 1800-tallets Danmark", in *Historisk tidsskrift*, 106, 1: 84, 86.

Howitz, Frantz Gotthard, 1824a, *Om Afsindighed og Tilregnelse, et Bidrag til Psychologien og Retslæren*, in *Juridisk Tidsskrift*, 8, 1: 1-117.

Howitz, Frantz Gotthard, 1824b, *Determinismen eller Hume imod Kant. Et philosophisk Forsvar for Afhandlingen om Afsindighed og Tilregnelse*, Seidelin, Kjøbenhavn.

Howitz, Frantz Gotthard, 1825, *Ultimatum, angaaende Determinismen og Etatsraad Ørsteds fortsatte Bemaerkninger om samme*, Seidelin, København.

Høffding, Harald, 1909b, "F.G. Howitz", in Id., *Danske Filosofer*, Gyldendalske Boghandel, Nordisk Forlag, København: 81-88.

Jacobsen, Anja Skaar, 2007, "Phrenology and Danish Romanticism", in Brain, Robert M., Robert S. Cohen, Ole Knudsen, eds. *Hans Christian Ørsted and the Romantic Legacy in Science. Ideas, Disciplines, Practices*, Series "Boston Studies in the Philoso-

phy of Science", Springer, Dordrecht: 55-74.

Kant, Immanuel, 1797, *Metaphysischen Anfangsgründe der Rechtslehre*, F. Nicolovius, Königsberg; Eng. tr. by John Ladd, *Metaphysical Elements of Justice*, Indianapolis/Cambridge 1999[2] (1965).

Kierkegaard, Søren, 1844, *Philosophiske Smuler eller en Smule Philosophi* (af Johannes Climacus), Reitzel, København (in *Søren Kierkegaards Skrifter* [*SKF*] bd. 4, 1997, Gads Forlag, København); Engl. tr. by Howard V. Hong and Edna H. Hong 1985, *Philosophical Fragments*, Princeton University Press, Princeton NJ.

Kiesewetter, Johann Gottfried, 1795, *Versuch einer fasslichen Darstellung der wichtigsten Wahrheiten der neueren Philosophie für Uneingeweihte*, Oehmigke, Berlin; Dan. tr. by Georg Johann Thomsen, *Den kritiske Philosophies vigtigste Sandheder for Uindviede*, Brünmich, Kjøbenhavn.

Koch, Carl Henrik, 1980, "Ørsted og striden om Viljens Frihed", in Tamm, D., (1980), *Anders Sandøe Ørsted 1778-1978. Foredrag i anledning af 200-året for Anders Sandøe Ørsteds fødsel*, Juristforbunden, København: 87-121.

Koch, Carl Henrik, 2003a, "Den danske kantianisme og dens kritikere", in Id., *Dansk oplysningsfilosofi 1700-1800*, Gyldendal, København: 34-44.

Koch, Carl Henrik, (2003b), "Anders Sandøe Ørsteds forsvar for Kant", in Id., *Dansk oplysningsfilosofi 1700-1800*, Gyldendal, København: 44-49.

Koch, Carl Henrik, 2003c, "Udsyn over det 19. århundredes danske filosofi", in Id., *Dansk oplysningsfilosofi 1700-1800*, Gyldendal, København: 357-360.

Koch, Carl Henrik, 2004, *Den danske idealisme, 1800-1880*, Gyldendal, København: 177-208.

Koch, Carl Henrik, 2008, *The Danish Debate about Free Will*, in J. Stewart (ed.), *Johan Ludvig Heiberg. Philosopher, Littérateur, Dramaturge, and Political Thinker*, Museum Tusculanum, Copenhagen: 11-36.

Lange, Frands Martin, 1822, *Forsete. Et Tidsskrift for Udvalg Af Høiesterets-Sager og for Andre Retsvidenskabelige Gjenstande*, nos. 15-16-17.

Lange, Frands Martin, 1835-1841, *Udvalg af danske og udenlandske Criminalsager og af mærkelige Forhandlinger om saadanne, samlet efter Retsacter, de bedste fremmede Værker og fra andre Kilder*, København.

Locke, John, 1690, *An Essay Concerning Human Understanding*, Thomas Bassett, London.

Michelsen, Knud, 1989, *Synålejomfruen og lægevidenskabens menneskeopfattelse*, Munksgaard, København.

Michelsen, Knud, 2008, "Howitzfejden", in S. Auken (ed.), *Dansk litteraturs historie. 1800-1870*, Gyldendalsk Bog, København, 2: 80-84.

Molbech, Christian, 1828-1833, *Dansk ordbog: indeholdende det danske sprogs stammeord, tilligemed afledede og sammensatte ord, efter den nuværende sprogbrug forklarede i deres forskiellige betydninger, og ved talemaader og exempler oplyste*, Gyldendalske Boghandling, Kjöbenhavn.

Nørregaard, Lauritz, 1784a, *Forelæsninger over den Danske og Norske Privat Ret*, Gyldendals Forlag, Kjøbenhavn.

Orfila, Mathieu, 1823, *Leçons de médecine légale*, Béchet jeune, Paris.

Ørsted, Aanders Sandøe, 1798, *Over Sammenhængen mellem Dydlærens og Retslærens Princip*, Seidelin, Kjøbenhavn.

Ørsted, Aanders Sandøe, 1815-1822, *Eunomia eller Samling af Afhandlinger, henhørende til Moralphilosophien, Statsphilosophien og den Dansk-Norske Lovkyndighed*, Seidelin, Københavnn.

Ørsted, Aanders Sandøe, 1824a, "Et Par Ord i Anledning af den foranstaaende Afhandling", in *Juridisk Tidsskrift*, 8, 1: 118-154.

Ørsted, Aanders Sandøe, 1824-1825, "Fortsatte Betragtninger over den menneskelige Frihed og Sædelærens Grundbegreb; i Anlendning af Hr. Professor Howitz's nye Skrift", in *Juridisk Tidsskrift*, 9, 2; 10, 1.

Otto, Carl, 1825, *La dispute en Danemarck sur le liberté morale*, in *Messager Français du Nord*, 1: 197-213.

Paul, Fritz, 1973, *Henrich Steffens. Naturphilosophie und Universalromantik*, Fink, München.

Pinel, Philippe, *Traité médico-philosophique sur l'aliénation mentale*, Brosson, Paris 1809[2].

Prichard, James Cowles, 1835, *A Treatise on Insanity and Other Disorders Affecting the Mind*, Sherwood, Gilbert and Piper, London 1835.

Quetelet, Lambert Adolphe Jacques, 1831, "Recherches sur le Penchant au Crime aux différents ages", in *Noveaux Mémoires de l'Académie*, 7, 1; Eng. tr. by Sylvester, Sawyer F., *Research on the Propensity for Crime at Different Ages*, Anderson Publishing Co., Cincinnati OH, 1984.

Radzinowicz, Leon, 1966, *Ideology and Crime. A Study of Crime in its Social and Historical Context*, Heinemann, London.

Skålevåg, Svein Atle, 2002, "Sykdom og tilregnelighet fra sakkyndighetens historie", *Tidskrift for Den norske legeforening*, 1: 65-68.

Skålevåg, Svein Atle, 2016, *Utilregnelighed. En historie om rett og medisin*, Pax Forlag, Oslo.

Steffens, Henrich, 1803, *Indledning til philosophiske Forelæsninger*, A. Seidelin, Kjøbenhavn; udg. af J. Kondrup, Det Danske Sprog- og Litteraturselskab, C.A. Reitzel, København 1996.

Stewart, Jon, 2003, *Kierkegaard and His Contemporaries. The Culture of Golden Age Denmark*, De Gruyter, Berlin-New York.

Stewart, Jon, 2007, *A History of Hegelianism in Golden Age Denmark*, Tome I: *The Heiberg Period: 1824-1836*, Søren Kierkegaard Research Centre - C.A. Reitzel, København, Chapter 2: *Heiberg and the Howitz Controversy: 1824*: 136-190.

Stewart, Jon, 2009, *Johan Ludvig Heiberg: Kierkegaard's Criticism of Hegel's Danish Apologist*, in J. Stewart (ed.), *Kierkegaard and his Danish Contemporaries*, Tome I:

Philosophy, Politics and Social Theory, Ashgate, Aldershot, Hampshire.

Stewart, Jon, 2014, *The Cultural Crisis of the Danish Golden Age: Heiberg, Martensen and Kierkegaard*, Museum Tusculanun Press – Søren Kierkegaard Research Center – University of Copenhagen, København: 85-88.

Thomsen, Oluf, 1924, *F.G. Howitz (1789-1826) og hans strid om "villens frihed"*, Levin & Munksgaard, København.

Waaben, Knud, 1997, *Retspsykiatri og strafferet i historiens lys*, Janssen-Cilag, København.

Waage, Olafur, 1867, *Strid om den menneskelige Villies Frihed og Sædelærens Grundlag*, in Id., *J.P. Mynster Og De Philosophiske Bevægelser Paa Hans Tid i Danmark*, C.A. Reitzel, København: 39-104.

On madness and ascribing responsibility

Frantz Gotthard Howitz
(1789-1826)

§. 7.

On Freedom according to the Kantian System. Objection against it based on the Conditions of Unfreedom

If I have otherwise correctly understood Kant's view, it can be summarized by the following 4 propositions:

1st Proposition. The will is the faculty of desire, the ground of which lies in reason. The will can be called neither free nor unfree, for it is practical reason itself.[1] [31]

2nd Proposition. a) The arbitrary will (*arbitrarium liberum*) is the ability to choose from opposite grounds or the ability to prefer the feeling of passion to a principle, or vice versa.[2] b) It is in this arbitrary will that freedom expresses itself as practical, but "this freedom is no *libertas indifferentiae*, for it does not exist in the ability to choose for or against the law" (Kant, l. c., p. XXVII)[3]. Freedom only goes in the one direction and is (positively considered) "the dependence of the arbitrary will on reason or the law of pure morality, but negatively considered, it is the independence of the arbitrary will from all sensuous

* "Om Afsindighed og Tilregnelse, et Bidrag til Psychologien og Retslæren", in *Juridisk Tidsskrift* 8, 1 (1824): 1-117. The numbers that appear in square brackets refer to the original pagination. The footnotes are Howitz's own. Editorial integrations appear in square brackets.

1 Kant's *Metaphys.*[*ischen*] *Anfangsgr.*[*ünde*] *der Rechtslehre*, [F. Nicolovius], Königsberg 1778 [1797]. *Introduction to Die Metaphysik der Sitten*, p. V and p. XXVII. As far as I know, this is the latest of Kant's writings about this subject and must therefore be regarded as containing his definitive view. On this see C.[arl] C.[hristian] E.[rhard] Schmid, *Wörterbuch zum leichteren Gebrauch der kantischen Schriften*, [Cröker], Jena [1786] 1795 [dritte vermehrte Ausgabe], p. 222.

2 P.[eter] E.[rasmus] Müller's, *Kristeligt Moralsystem*, [Brummer], Kjøbenhavn 1808. §. 35. Cf. [Johann Gottfried] Kiesewetter, *The Most Important Truths of the Kantian Philosophy for the Uninitiated* [*Versuch einer fasslichen Darstellung der wichtigsten Wahrheiten der neueren Philosophie für Uneingeweihte*, Oehmigke, Berlin 1795], translated into Danish by [Georg] J.[ohan] Thomsen, [*Den kritiske Philosophies vigtigste Sandheder for Uindviede*, Brünmich], Kjøbenhavn 1797, p. 156. [Cf. Kant, AA VI, 226; Engl. tr. by John Ladd, *Metaphysical Elements of Justice*, Indianapolis/Cambridge 1999², p. 19].

3 [Kant, AA VI: 226; Eng. tr.: 19].

ISSN (print) 2281-8618-ETS DOI: 10.4454/philinq.v8i1.284

motivations, as necessary grounds for action" (Schmid, l. c., p. 223).

The mere *arbitrarium liberum* is not enough to give us the name of moral beings, since the ground of possibility of all morality is contained in freedom (Kiesewetter, p. 154).[4]

[32]

3rd Proposition. Freedom in the absolute or transcendental sense is, when regarded negatively, the independence of the will from everything empirical, from all the laws of natural necessity. But, when viewed positively, it is the absolute spontaneity or ability to begin a new series of causes.

4th Proposition. Freedom is not an object of theoretical knowledge; it is a mere regulative principle grounded in our supersensible nature, and it proves its reality by the fact that the moral laws make themselves known to us as categorical imperatives.

To attain for this summary presentation the appropriate degree of precision, we must add the following 5 corollaries:

1st Corollary. Human freedom is not merely something ideal but rather a real property. People do not merely have a predisposition to be free, but they really are.

2nd Corollary. Human beings can act immorally and yet possess absolute freedom. The action is then said to be done with freedom,

3rd Corollary. The senses cannot be a necessary reason for an immoral decision. The senses can only "affect" but not determine; they can give rise to an occasion and a temptation, but insofar as the human being is free, it is left to him if he will overpower the temptation or give in to it. Moral evil is therefore not in the senses themselves, [33] but in the fact that the will (*sensu latiori?*) with freedom submits to them. (*Eunomia*, vol. 2, p. 120).[5]

4th Corollary. a. Even in the greatest temptation and passion, the human being must still be considered free, that is, it must be possible for him to determine himself in agreement with the moral law. For he has consciousness of this law, and he criticizes himself for the action, and it is imputed to him both morally and legally.

b. When one talks about *degrees of freedom*, by this is not meant any increase or decrease in this property itself, but an increase or decrease in the affections of the senses that the will has to overcome.

5th Corollary. The origin of moral evil is incomprehensible, and its existence

[4] For this author, practical reason and liberum arbitrium are synonyms. But this is not in agreement with either the aforementioned work by Kant or with Schmidt.

[5] [The reference is to Anders Sandøe Ørsted, *Eunomia eller Samling af Afhandlinger, henhørende til Moralphilosophien, Statsphilosophien og den Dansk-Norske Lovkyndighed*, Seidelin, København, A.S., 1815-1822].

in the world inexplicable. (Kant, l. c., p. XXVIII).[6] There will always be an x that must, so to speak, be the complement to the affection of the senses and makes them the basis of determination. This x is just as incomprehensible regardless of whether one expresses it, as in Corollary 3, as a voluntary submission or explains it as the inactivity in practical reason such that the human being fails to assert the freedom he possesses.

Such is Kant's doctrine of freedom. It cannot escape the initiate's attention that Kant in this appears to be the complete antipode of Augustine. For Augustine says, "People only have the freedom to sin since the freedom not to sin has been taken from them with the fall of Adam" ([Wilhelm Gottlieb] Tennemann, *Grundriß der* [34] *Geschichte der Philosophie*, [Barth], Leipzig 1812, p. 161). Kant says that human freedom only leads to good and that it is the dependence of the arbitrary will on the pure moral law (Proposition 2). Augustine cannot, after what has been mentioned, explain the good works of human beings without assuming the influence of grace. Kant, according to his theory, cannot explain the immoral deeds of human beings (Corollary 5) or the origin of moral evil. In general, Kant is close to Pelagianism. The very deduction of his conception of freedom from the existence of the moral law (Proposition 4) will be recognized in the following sentence of Celestius, a disciple of Pelagius: "If now man should be without sin, then he can be without sin, and if he cannot, then it could likewise not be obligatory" (Holberg, *Church History*, Copenhagen 1740, 1st part: 220).[7] But Kant is an Ultra-Pelagian in so far as his conception of freedom is not a freedom of choice (a *libertas indifferentiae*) between good and evil (Proposition 2) but the freedom of virtue, or, in other words, a freedom that not only presupposes the possibility that a human being can act morally well or rationally, but the necessity that he do so, as often as this freedom is really expressed; for this can only be the case in accordance with its so-called autonomy (self-legislation), which is infallible.

There is an apparent similarity between Kant's definition of "freedom" and Spinoza's: "*humanam potentiam in moderandis et coercendis affecti-* [35] *bus servitutem voco, homo enim affectibus obnoxius sui juris non est sed fortunae*" (*Ethics*, pars IV in the *Preface*)[8] and *ibid*., Proposition 68: "*Illum libe-*

6 [AA VI: 226; Eng. tr.: 19].

7 [Ludvig Holberg, *Almindelig Kirke-historie fra Christendommens første Begyndelse til Lutheri Reformation, med nogle Anmærkninger over de udi historien omtalte Cyclis og Aars-Beregninger*, Høpffner, Kjøbenhavn 1740[2] (1738)].

8 ["I assign the term 'bondage' to man's lack of power to control and check the emotions. For a man at the mercy of his emotions is not his own master but is subject to fortune", Eng. tr. by Samuel Shirley, in B. Spinoza, *Complete Works*, ed. by Michael L. Morgan, Hackett Publishing Company, Indianapolis-Cambridge 2002: 320].

rum esse dixi, qvi sola ducitur ratione".[9]

But Spinoza's conception of freedom is an ideal (*De jure natura*, l. c. XVIII)[10], to which human beings can only approach and which, in general, they are very far from because their reasoning is limited and their emotions predominant (*Ethics*, l. c., [Pars IV] Proposition 37, Scholium II). By contrast, Kant's conception of freedom is a property with which he assumes the human beings in general are endowed, indeed even the most vicious, light-minded and irrational, as long as they have not gone over into a condition of total unfreedom due to madness. Spinoza's conception of freedom has as its basis reason as a whole and a reason that aims at self-preservation and true bliss, a reason that has a common root and common purpose with the senses (*De jur.*[*e*] *nat.*[*urali*], § 5) and can oppose the latter, as often as their promptings would lead away from the goal by the further consequences of the action. By contrast, Kant's conception of freedom has as its basis the [36] mere practical reason or the pure moral law, but not theoretical reason or the faculty of inference, which allows us to foresee the consequences of action, and this practical reason has nothing in common with the desire for happiness or with the senses. A foolish act cannot be called free according to Spinoza, but it can according to Kant. Finally, the conception of freedom in Spinoza is consistent with the doctrine of necessity and the knowledge of the eternal laws of nature and an eternal series of causes, which also includes all human thought and action, "*agendi necessitate non tollit sed ponit libertas* (*De jure naturali*, [Cap. II] § XI)[11] *nihil namqve homo, seu ratione seu sola cupiditate ductus agit nisi secundum leges & regulas naturae* (*ibid.* § V)[12]. Kant's conception of freedom, by contrast, makes man a being independent of natural causes, who can intervene in them at any moment with his own arbitrary self-determination, and whose existence here in the world must therefore be regarded as an "*imperium in imperio*", more suited to disturb than to confirm the natural laws to which all other beings are subject (*Ethics*, p. III.

9 [*Ibid.*: 355: "A free man is he who is guided solely by reason"].

10 Likewise in his treatise *De jure natura* in [Chap. II], § XI. "Imo qvia humana potentia non tam ex Corporis robore qvam ex mentis fortitudine aestimanda est, hinc seqvitur, illos maxime sui juris esse, qvi maxime ratione pollent, qviqve maxime eadem ducantur, atqve adeo hominem eatenus liberum omnino voco, qvatenus ratione ducitur" [Eng. tr. by Samuel Shirley, *Political Treatise*, in B. Spinoza, *Complete Works*, op. cit.: 686: "Human power should be assessed by strength of mind rather than robustness of body, it follows that those in whom reason is most powerful and who are most guided thereby are most fully in control of their own right. So I call a man altogether free insofar as he is guided by reason because it is to that extent that he is determined to action by these causes"].

11 [*Ibid.*: "Freedom does not remove the necessity of action, but imposes it"].

12 [*Ibid.*: 683-684: "Whether a man is led by reason or solely by desire, he does nothing that is not in accordance with the laws and rules of Nature"].

in the *Preface*)[13]. Spinoza's conception of freedom is thus ideal, but thereby natural. Kant's conception is real but supernatural. Spinoza's conception of freedom is rationality by means of the drive to happiness, whereas Kant's is morality by means of categorical imperatives. Spinoza's doctrine is affirmed by experience and can be understood by people, whereas Kant's doctrine is contrary to experience and presupposes the incomprehensible. [37]

One remark can be made against both of these philosophers' definitions of freedom: the fact that they have given this word a somewhat divergent meaning from that of ordinary linguistic usage. In everyday language, freedom is opposed to coercion or limitation, but it is not opposed to prosperity, simplicity or immorality. A free man is the one who can "do what he wants", act according to his own wishes, choose according to goodwill, and follow the promptings of his nature, without being dependent on foreign laws or foreign wills. But it is not asked whether these wishes are morally good or bad, whether this goodwill is rational or irrational, whether these promptings are virtuous or sensuous; in short, there is the same difference between Spinoza's or Kant's views and the ordinary interpretation of the word "freedom" as there is between the two expressions "to be his own master" and "to be master of oneself".

However, since no one is more of his own master than the one who is also master of himself, the use of the word "freedom" can in a sense be defended, but it easily gives rise to misunderstanding. Thus, it is no rarity in Kant's writings to find the word "freedom" and especially the adjective "free" sometimes used in the strict Kantian sense and sometimes as synonymous with arbitrariness or freedom in the common understanding. One will find this misuse right in the first sentence by Kant, "The will can be called neither free nor unfree" (Proposition 1). For the will, which itself is absolute spontaneity, [38] must indeed be free, according to Kant's own definition. The thing is that "free" here is opposed to "unfree", that is, involuntary.[14]

Now we come to an argument against Kant's doctrine of freedom, which stands in the closest connection with the main subject of this treatise, as it is taken from the view of the condition of unfreedom. Indeed, it will be appreciated that the origin and existence of unfreedom, which Kant cannot deny, is just as incomprehensible and inexplicable according to his system as moral

[13] [B. Spinoza, *Ethics*, op. cit.: 277: "Most of those who have written about the emotions and human conduct seem to be dealing not with natural phenomena that follow the common laws of Nature but with phenomena outside Nature. They appear to go so far as to conceive man in Nature as a kingdom within a kingdom"].

[14] A similar misuse is with the word "will", which sometimes is used in the ordinal sense and sometimes in agreement with Kant's definition. Such results are inevitable when one removes words from their usual usage.

evil itself, and the closely related interrelationship of the two, the parallelism of the physical, intellectual and moral pathology, their common root in man's sensuous nature (*sensu latiori*), certainly contributes not a little to weaken the Kantian principle "that man as a moral being is independent of everything empirical or possesses the property of being determined by the senses not by necessity (the 3rd Proposition, and the 3rd Corollary).[15]

[39]

This sentence contains in and of itself something offensive to anyone who is familiar with the physical side of man, and I dare say with science at all. For man is for him not two individuals but one; therefore, no single part is independent of the whole. Man is not *animal* and *ratio* but an *animal rationale*, not body and soul but a besouled body, and whatever be the origin of these different properties, the human being here in the world consists in and by their fusion to a whole. It is one thing by abstraction to distinguish between the intellectual (thinking) and the physical (extended) properties, and another to make them independent of each other. A physiologist can distinguish between the animal and the vegetative elements in man, between movement and nutrition, between nerves and muscles, on one side, and blood and the lymph system, on the other side, but he says an absurdity when he claims them to be independent of each other: "*omnia in homine se habent ad instar circuli et ubi* [40] *qveris initium invenies finem et ubi qveris finem invenie initium*".[16]

The experience of millennia has taught that the human race, as long as it has existed on earth, has always been dependent on the senses in the moral understanding (that is, desires, inclinations, passions). This experience has taught that not only are most people determined by these (which is irrefutable) but that even the best ones have been drawn to acts contrary to the precepts of reason and moral law. Thus, it still is now, and, presumably, it will remain this way for as long as the world exists.

The fact that Kant's doctrine of freedom is in conflict with this old and common experience, which it can only admit as being the effect of an incomprehensible cause (cf. Corollary 5), is certainly not able to make it probable. However, one is accustomed to regard moral evil as a mystery, to whose dis-

[15] I refer especially to this proposition along with the one mentioned under the 4th Corollary. If a defender of Kant were to come forth and prove that I had completely misunderstood Kant's doctrine, then I would ask him whether Kant did not assume "that the human being in the great temptation *can* resist when he want, and that such a will is always possible, unless a total unfreedom is present". If he were to grant me this point (which he doubtless must), then I do not see that another misunderstanding of the system could have special influence on the objections mentioned hereafter, which are grounded in the constant dependence of the will on the physical nature of the human being.

[16] [Hippocrates, *De locis in homine*, 2, VI, 276 L].

solution man in an emergency fetches the devil from a background of ideas, where he still stays.

But the dependence of the will upon the *physical senses* of man, that is, his organization, his physical development, and his state of health, is a just as old and undeniable experience as the one just mentioned, and the contradiction in which it stands to Kant's doctrine of freedom seems the stronger, the more unmistakable and obvious the necessity is with which the will [41] (especially in the so-called condition of unfreedom) is determined by physical causes.

The moral nature of human beings, just as the intellectual nature, apparently depends on the brain's organization and development (childhood), on its diverse vitality (sleep, dreams, intoxication), on its stage of decline (old age) and on its diseases (madness). "One has" says Parry (l. c. § 770) "seen a random blow to the head determine the best human principles, and transform a pious Christian into a drunk and an irredeemable criminal" (cf. § 3. Note quoted).

One sees again the moral character of human beings as determined by innate temperament and hereditary drives.[17] One sees it change according to climate and diet, and this even becomes apparent throughout entire nations (northerners and southerners, plant-eaters, fish-eaters, meat-eaters). One finds it different by gender and age and state of health. It is one thing in a man another in a woman, one thing in a younger man and another in an old man, one thing in [42] the feeling of strong health, and another in sickness, weakening, pains, hunger, insomnia, etc.

In addition to the immediate dependence of morality on the physical elements, there is also given an indirect dependence through the intellectual organs and their connection with the external senses. Any single *actus* of "faculty of desire in accordance with *concepts*"[18] presupposes the association of ideas, memory, or impressions of the senses that all depend on natural causes, and practical reason is not without the necessity connection with the theoretical element, the degree and activity of which again rises and falls with the angle of the face and the height of the forehead. The New Hollander[19] who belongs to a human race that stands only a little above the beast in his understanding is also extremely lazy, cowardly and lustful. Everywhere it seems the wild man

[17] "Fortes generantur fortibus & bonis, nec imbellem parit aqvila columbam". If this sentence is not literally true, then it is valid in most cases. Not all siblings look alike, but most do; not all children look like their mother or father, but most do. "Non parum felices bene nati!"

[18] [Kant, AA VI, 213: *Begehrungsvermögen nach Begriffen*].

[19] [At the time Europeans used the term "New Hollanders" to refer to the aboriginals of Tasmania and later of the whole of Australia. The name derives from "Hollandia Nova", the name that the Dutch navigator Abel Tasman gave to Australia in 1644. The term remained in use at least until the mid-nineteenth century].

possesses only the virtue of temperament: he is good where he loves, but evil where he hates. Everywhere, the child is self-serving and guided by the senses to a great extent, the old person in general also. In addition, the old person becomes twisted, biased and unfair in relation to how the physical forces and with them the memory and the other heart functions diminish. Sometimes he is even well aware of what the moral law offers. So if he has freedom, why doesn't he use it as well as before?

Undoubtedly, it will be said that all the different circumstances listed here must be attributed either to perfect states of unfreedom such as intoxication and madness or [43] to overpowering affections of the senses, such as temperament, character and temporary moods; and if one wants to refer to Socrates, who, though born with a grossly sensual temperament, became wise and meek, or Xenocrates, who remained cold beside Phryne[20], although he had recently drunk much wine. But is it possible for all people to act like these two philosophers? Was Socrates' mastery over his sensuality not more the fruit of a rationality, fought for and acquired, than of an original freedom (Corollary 1), and would he then not be an exception to the rule, and more to be regarded as an ideal than as a true copy of the existing race? As for the wise Xenocrates, his freedom is so little human that I do not even know if it was worth possessing. However, we would like for a moment to let it remain undecided whether the differences in temperament and character have a necessary effect or not, and turn to the perfect states of unfreedom.

Unfreedom is the lack of freedom, therefore, it is the lack of those attributes that constitute the essence of freedom (Propositions 2 and 3). But, I ask, how does this lack arise? How it is conceivable that an attribute, whose basis is absolute spontaneity and independence from everything empirical and all natural necessity, can disappear because the stomach comes into contact with the spirit of wine, or because the head [44] is exposed to the sun's rays, as King Charles 6th's story[21] teaches, or the example of the Abderites[22], who became insane because, during the performance of Euripides' *Andromeda*, they forgot to cover their heads against the sun? For what reason does this supersensible

20 [Phryne was a famous ancient Greek courtesan (hetaira) of the 4th century B.C. Diogenes Laertius tells that she tempted Xenocrates in vain to enter her bed: but she later reported that it was like sleeping with a statue.]

21 [Charles VI of Valois called "the Mad", king of France from 1380 to 1422. It is said he had his first outbreak of madness after a prolonged exposure to the sun on an August day].

22 [Cf. also *Der Streit der Fakultäten* (1798), AA VII, p. 82, where there is mention of the frequent mood swings of the Abderites. "Abderitism" is defined as the philosophy of history according to which mankind will proceed in the moral sphere according to a constant and changing ascent and an equally frequent and profound relapse, with the result of finally remaining in a state of immutability.]

sovereignty resume as soon as the intoxication has been slept off, or the blood in the brain diminished by leeches and Spanish flies, and what resolution can be found to the mystery that it in cases of partial madness is partly present and active, and partly not?

All this, so it is stated, is inexplicable and rests on the eternally incomprehensible connection between soul and body – another inexplicability of Kant's doctrine and another struggle against experience where a Spinozist does not find the least difficulty.

But does this incomprehensible lack of freedom also abolish the arbitrary will (*arbitrium liberum*)? One should think not! One would think that the moral law might change from master to counselor without therefore disappearing, and that its precepts could sink from categorical imperatives to simple motivations that could be overcome by opposing motivations or could be victorious over them depending on their relative weight, that is, in accordance with how the human being judges that his well-being is best advanced by this or that.

However, this kind of arbitrary will, no matter how likely it is made by experience, would be totally inconsistent with Kant's principles. According to these, [45] no human or moral arbitrariness can be conceived without freedom; for practical reason, the pure will, must possess absolute spontaneity and mastery in the arbitrary will, as often as it appears, or it must disappear altogether; it must be *aut Caesar aut nihil*. The restriction of liberty in man lies in the temptation of moral evil, (Corollaries 3, 4 and 5), but it is not in the decline of freedom itself, not in the impudence of practical reason or in the necessary advantage of the senses over it; for it follows from the nature of Kantian freedom that the pure will can never be overcome with necessity. It is a giant whom the senses (Hedone) may well entice to moral evil, but in an open battle she must always be inferior to them. Unfortunately, this giant is also inclined to leave his post; there arises then a kind of interregnum (unfreedom), and in this the senses play the master. According to this, one must therefore state that freedom is either completely present or not at all, that it is *aut Caesar aut nihil*, and that the not-moral is either an effect of a morally evil will or of a condition of total unfreedom, without one being able to discern the further connection between the two than some gradual transition from the one to the other. I have gone into great detail in presenting this Kantian proposition partly because it [46] belongs to the more hidden part of this doctrine and partly because I consider it both false and damaging in its application. My objections are as follows:

1) We saw above that practical reason was necessarily dependent on theoretical reason. In the child, in the elderly person, in the poorly endowed human races, in the uncultivated savage, we saw that morality was modified by intellectual development. But this has degrees, and these degrees are due to

natural causes (congenital disposition, upbringing, chance circumstances), and therefore practical reason and with it freedom are indirectly subjected to the same natural causes and the same gradual development and decline. So there is a gradual transition from freedom to unfreedom.

2) The Kantians admit that freedom is thus dependent on natural causes and that it, albeit in an incomprehensible way, can be suspended by a tablespoon of blood in the brain beyond the usual quantity (intoxication, delirium, etc.). But if this is so, what prevents them from assuming that the same freedom can, due to similar reasons, be limited with necessity so that it ceases to be Caesar without just becoming *nihil*? And, if freedom and reason can thus be suspended or limited by the blood that the wine, solitude or feverish paroxysm moves to the head, then why can't the same effect arise in the human being who is burning with anger, revenge, shame or sex drive?

[47]

3) Experience testifies in the most determinate manner to boundaries between freedom and unfreedom and opposes any attempt to limit the two. If there is any sharp boundary between an unfree child and a child deserving punishment, between sluggish and weak old people, between simplicity and silliness, between depression and melancholy, between hot temper and fury, between enthusiasm and ecstasy, between exaltation and intoxication, between resistible and irresistible bodily drives (hunger, sex drive, sleepiness, etc.), between the states when the voice of reason and morality can still be heard and those in which is silent. Don't let jurisprudence have to set arbitrary limits here; but should morality, should philosophy follow its example, shall jurisprudence itself affirm its approach as perfect, its judgment as infallible, when it is only a necessary consequence of the hitherto acknowledged limitations of human discernment?

Thus, I have endeavored to prove the correctness of the three propositions set out in § 7, which prevented me from constructing the concept of madness on the Kantian doctrine of freedom; but before I leave this subject, allow me yet to make a comment, without which any objections to Kant's doctrine would be powerless and strand on the Kantians' *a priori* conviction of the truth of their case.

[48]

According to Kant's 5 propositions, the assumption of freedom in the often discussed sense is an article of faith that no objection can weaken, and this article of faith is grounded in humanity's awareness of the moral law as an unconditional imperative. But is the matter with this imperative really as Kant claims? Is there an original *should*, implanted in all human beings as a rule, without condition, without proof, without why? Locke, the astute and so

God-fearing Locke, assumed no such moral principle, *qvasi coelitus in mente descripta*; he contested its existence by noting how the customs and moral concepts of various nations differ from each other and how the past was often contradictory to the present. He assumed that virtue was widely esteemed because it was generally beneficial and that there was no moral rule that needed proof.[23]

Nor does it seem to me to be called into question that morality is grounded in human coexistence and mutual social relations, that it first with these is evoked, and that the categorical imperatives would be utterly silent if man still lived *in statu solitudinis*. Only the principle "promote your happiness" would then be dominant,[24] and all so-called [49] duties to oneself would be attributed thereto, but such a person would never dream of maxims becoming the object of legislation.

However, such cold maxims were never the language of nature. Only an ivory tower scholar could regard them as such. Nature formed our hearts for sympathy, compassion and benevolence for beings who resemble us, who have with us a common origin, a common destiny, a common fear and hope; it made us increase and double our benefits by communication and society with these beings, and it put in us the desire to find our loving feelings answered by reciprocal love and recognition, our self-esteem reinforced by their approval and admiration. These feelings and the soon-gained experience of their beneficial influence were what created virtue and made it gracious and paired it with our sense of beauty and set it as a goal which human beings might be quite happy to strive for. But what was egoistic was hated, as unsocial and cruel. Thus, wisdom united with the sociable drives to create the morality.

[50]

Instead of this view, which makes virtue natural and related to other human desires, about which one can therefore say as about Socrates' doctrine that it lures morality from heaven and brings it into the heart of man[25], instead of that which gives us, as if a contradiction, a morality of categorical imperatives; a despotic rule without a "why", preached to us as if by inspiration or more correctly as if by the sealed orders that are given to expeditions on the ocean to be

[23] *Philos. Essay in Human Understanding* [*An Essay Concerning Human Understanding*, Thomas Bassett, London 1690], Lib. I, cap. II.

[24] Even the striving for perfection according to the development of the intellect or physical powers would not be present in this condition without their being brought forth by the desire which man, just like the animal, feels with activity, by overcoming obstacles, and seeing the fruit of his efforts, while the opposite feeling oppresses him. The actual feeling of honor arose first when the human being compared himself with his others of his kind.

[25] [The reference is probably to Cicero's famous statement, "Socrates autem primus philosophiam devocavit e coelo, et in urbibus collocavit, et in domos introduxit, et coegit de vita, et moribus rebusque bonis, et malis", cf. *Tusculanae Disputationes*, V 4, 10-11].

opened at a certain degree of longitude; a cold instrument that we, despite the resistance it finds in our dearest wishes, do not merely obey unconditionally but even gladly invest our freedom to obey.

It at least was an implanted sense of feeling of well-being for virtue and displeasure with vice, he believed, a kind of moral instinct analogous to our feeling for beauty or to the natural sympathy we call humanity; but no! Since, according to Kant, all feeling of desire and lack of desire is grounded in sensing being, in our lower desires, it would be unworthy to give the moral law an origin of this kind. This must be a rule of reason, and the moral feeling can only be something derived, something mixed. Nor can this rule be regarded as the advice of reason or guidance to happiness in society, based on the recognition of the beneficial consequences of virtue; it must be, according to Kant, a command, and a categorical command, that is, a command without ground.

It is then no longer virtue's own beauty, [51] not the pursuit of the height of greatness of the human soul and independence, not the desire to deserve the approval and love of man, not the joy of spreading happiness and satisfaction among my fellows, not the hope of the growth of the entire race in everything good and beautiful, and of an increasing rationality and happiness on earth, and of the gratifying consciousness of having contributed to it. It is nothing of all this that will make me noble and virtuous and sacrificing. It is no longer the virtue and fame of a Themistocles that will incite a Cimon, not the first example of Brutus that will raise a Regulus or Cato. Hereafter we are supposed to have Cimons and Reguli and Catos through the choice of maxims that could become common law. And precisely what makes virtue most gracious in human beings and gives it its greatest glory, I mean its origin in the feelings of a warm and benevolent heart, precisely this detracts from its worth for the Kantian, who places such virtue of temperament far below that which arises from that unconditional obedience to reason's *You Should*.

Kant's moral law, on which his entire doctrine of freedom is grounded, is little probable and natural, so cold and lacking in beauty, so far from the ancient philosophers' elevating and inspiring images of virtue and virtue's reward.

"*Les hommes*", I say with Rousseau, "*m'eussent* [52] *jamais été que des monstres, s'il a nature ne leur éut donné la pitié á l'appui de la raison*".[26]

Translation from Danish by Jon B. Stewart

[26] [Cf. Rousseau, *Discours sur l'origine et les fondements de l'inégalité parmi les homme*, Marc Michel Rey, Amsterdam 1775: 71; Eng. tr. by Donald A. Cress, Hackett Publishing Company, Indianapolis-Cambridge 1992: 37: "Men would never have been anything but monsters, if nature had not given them pity to aid their reason", in *Discourse on the Origin of Inequality*].

Citation Style

Philosophical Inquiries adopts a variation of Harvard author-date citation style.

Substantive notes are numbered, references are parenthetical (in-text, in round brackets), bibliography is compulsory. Compared to Harvad style, *philinq* uses a simplified punctuation and a "European" style with regard to Publisher-City order.

Quotations and punctuation

Quotation marks: double inverted commas. Use single inverted commas in nested quotations only.

Terminal punctuation and superscripts ought to be *outside* quotation marks.

Superscripts ought to be *outside* quotation marks and *outside* terminal punctuation.

In-text references ought to follow quotation marks.

Quoted question marks are of course considered part of the quotation.

Examples:

"quoted text".[superscript]

"quoted text"[superscript, if applicable] (reference).

"Is this a question?" (reference).

Use of capital letters

Essay title: Very first letter only.

References / book title: current English rules apply.

References / article title: Very first letter only.

References / translations: Very first letter only; keep capital letter in language abbreviation (e.g. "Eng. tr.").

ISSN (print) 2281-8618-ETS

In-text references

Books, essays, published/unpublished conference papers, etc… are all referred to through Author and year.

Commas are used to separate different (multiple) authors. Colons are used instead of "p." or "pp.". Semicolons are used to distinguish different works and the original edition from a translation in use (if applicable). Letters are used to distinguish different works written by the same author in the same year.

Single author:

> Sophisticated searching techniques are important in finding information (Rossi 1994).
>
> Rossi (1994) claimed that sophisticated searching…
>
> Rossi states that "sophisticated searching techniques are important in finding information" (1994: 33).

Multiple authors: add *et al.* (always *italicised*)

> Sophisticated searching techniques are important in finding information (Rossi *et al.* 1994).

Translation:

> Rossi states that "sophisticated searching techniques are important in finding information" (1994: 33; Eng. tr. 1998: 28).
>
> *abridged form*
>
> Rossi states that "sophisticated searching techniques are important in finding information" (1994: Eng. tr. 28).

Multiple works:

> Sophisticated searching techniques are important in finding information (Rossi 1994; 1996).
>
> Sophisticated searching techniques are important in finding information (Rossi 1994; Bianchi 1999).
>
> Sophisticated searching techniques are important in finding information (Rossi 1994a; 1994b).

Avoid redundancies:

> Rossi has been working on this subject for almost a decade. In his essay on searching techniques (1994), he states that they are "extremely important in finding information" (33).
>
> Rossi has been working on this subject for almost a decade. In his essay on searching techniques (1994), he states that they are "extremely important in finding information" (Eng. tr.: 28).

Omonimies: use initial(s)

Sophisticated searching techniques are important in finding information (Rossi G. 1994).

Edited books

In their work on searching techniques (Rossi 1994), the Italian team on Information Engineering…

References

References are ordered alphabetically by last name. First personal names can be either in short form or extended, but please make a choice and don't mix. Multiple works from the same author will be ordered by year, most recent to oldest.

Elements of the citation

- Books

  Author of book – last name, first name abridged, Year of publication, ed(s). if applicable, *Title of book – italicised*, Translator if applicable, Edition if applicable, Publisher, Place of publication.

  Rossi G., 1994, ed., *Information Engineering*, ETS, Pisa.

MULTIPLE AUTHORS (UP TO THREE)

First Author of the book – last name first name abridged, Other Author(s) last name first name abridged, Year of publication, ed(s). if applicable, *Title of book – italicised*, Edition if applicable, Publisher, Place of publication.

Rossi G., Bianchi F., Verde G., *Information Engineering*, ETS, Pisa.

MULTIPLE AUTHORS (MORE THAN THREE): YOU CAN USE ET AL.

Rossi G. *et al.*, *Information Engineering*, ETS, Pisa.

TRANSLATIONS (WHEN QUOTED)

Ricœur, P., 1983, *Temps et récit*, Seuil, Paris; Eng. tr. by D. Pellauer 1990, *Time and Narrative*, University of Chicago Press, Chicago.

PLEASE NOTE THAT ABBREVIATIONS FOR LANGUAGES KEEP THEIR CAPITAL LETTER (Eng., Fr., Ger., It., etc…).

- Chapter(s) in a book

  Author(s) of chapter – last name, first name, Year of publication, "Title of chapter – in double quotation marks", in Editor(s) – family name, first name ed(s)., *Title of book – italicised*, Edition, Publisher, Place of publication: Page numbers.

NOTE: when multiple chapters from an edited book are quoted, reference style should reflect non-redundant criteria, thus including a separate reference for the edited book and abridging chapter reference as follows:

Rossi G., Bianchi F., 2013, "Information Engineering", in Rossi 2013: 9-24.

[i.e.: Author(s) of chapter – last name first name, Year of publication, "Title of chapter – in double quotation marks", in Editor(s) Last name Year of publication if not redundant: Page numbers.]

- Journal article
  Author(s) – last name first name, Year of publication, "Title of chapter – in double quotation marks", in *Title of journal – italicised*, Volume, Issue or number: Page number(s).

  Bhabha H.K., 1985, "Signs Taken for Wonders: Questions of Ambivalence and Authority under a Tree Outside Delhi, May 1817", in *Critical Inquiry*, 12, 1: 144-165.

NOTE: With regard to journals, the wide variety of choices and practices may pose several dilemmas. Few possibilities are hereby explored:

- Yearly journals: use volume number only.
- Double issues: use hyphen between each number.
- Special issues: cite extended special issue title if applicable, either add "s" before page numbers, either follow what is written in the front page:
  - Harper, P.B., A. McClintock, J.E. Muñoz and T. Rosen, 1997, eds., *Social Text*, 52-53, *Queer Transexions of Race, Nation, and Gender.*
  - …*Critical Inquiry*, 12: s1-s18.
  - …*Critical Inquiry*, 12, 1 bis: 1-18.
  - …*Critical Inquiry*, 12, 1, Spanish edition: 1-18.

- Web page
  Author(s) of page last name first name, OR organisation name Year (page created or revised), *Title of page – italicised*, Publisher (if applicable), Place of publication (if applicable), visited dd month yyyy, URL.

- Conference paper

PUBLISHED

"Chapter in a book" criteria apply. *Title of proceedings – italicised* – may include place and date

UNPUBLISHED

Author(s) of chapter last name first name, Year of publication, "Title of paper – in double quotation marks", paper presented to Title of congress meeting should include place and date.

Stephen T. Asma and Rami Gabriel
The Emotional Mind. The affective roots of culture and cognition
Harvard University Press 2019, pp. 448

Elena Pagni

*The Emotional Mind**, hereinafter called *tEM*, is an important book, not only because it offers a synthetic and syncretic sketch of theoretical perspectives, experimental data and empirical evidences recently achieved in the "mindbrain sciences", but also because it throws a new look and a greater awareness concerning the evolution of the mindbrain, focusing on questions and themes that have marked the historical path and development of philosophy, psychology and biology.

tEM also offers an original research program to understand the nature of emotions, their non-linguistic structures and cognitive value; in other words, the way in which, from an evolutionary point of view, the system of emotions is located at the interface between physiology and cognition, also showing how the development of language and concepts, and in general, of the human cultural evolution, is the complex result of a strengthening/reinforcement and enhancement of the biological value of brain-based social emotions, emerged together with ecological pressures (behaviors), lastly refined and extended over time by the frontal and neocortical abilities of *sapiens* cognition (219). "Representational abilities were decoupled from perceptual tasks, expanding possibilities for simulation and executive cognition abilities" (153). All this leads the authors to throw doubt on the assumptions of the rational choice theory (rational action theory) and, conversely, to rethink the evolution of human reason as based on their view of affective dominance (218).

I would like to stress the richness of this work and the correlative remarkable research program outlined; nevertheless, especially with regard to Descartes, I would have expected more interest towards his book *The Passions of the Soul* (1649) that represents, to my opinion, the great *ante-litteram* precursor of this approach to the emotional mind. Certainly, Descartes's conceptualization of emotions, in 1649, is comprehensively out of any evolutionary epistemological context. However, I would like to emphasize its importance for the cognitive value associated with emotions and the "plasticity" (in this time thought as

ISSN (print) 2281-8618-ETS

change of habit) assigned to the functioning of the mind in the activity of coding and recoding, and then interpreting, the emotional data.

The "affective turn" framework defended in *tEM* pursues the main objective to shed a new light on the evolution of the mind and the biological roots of emotions, "deeply rooted in what we know about the brain as a biological reality" (2); the assumption, that is, that the development of the emotional system is to be considered pivotal to understand the evolution of the human (and nonhuman) mind, in view of a new archaeological insight concerning the processes that led to the sedimentation and stratification of the different functional layers, ultimately overturning the assumed hierarchy of values.

> Our approach in this book is to show how the lowest layers of mind permeate, infiltrate, and animate the higher layers. The evolution of mind is the developmental story of how these layers emerged and acted as feedback loops on each other (10)

What is even more important to foreshadow, as Asma and Gabriel point out, is that such feedback is not a brain process, but "an embodied, enactive, embedded and socio-cultural process" (10). Hereinafter clearly referred as the formulation of a bottom-up cognitive model (that relies upon evolutionarily earlier manifestations of mind and social intelligence), as distinct from a top-down model of the mind (76).

It appears immediately clear, therefore, the attempt by the authors to situate themselves in opposition to the computational mindbrain metaphor and against social constructionists, by proposing a biological view of the roots of the mental that may assign an eminently adaptative value to the physiological-perceptive, emotional and cognitive components of the mental. "Affects are adaptations […] in two ways : phylogenetically (as evolved dispositions) and behaviorally (as real-time responses that may be a product of genes, learning, or cultural shaping). Affects are adaptations to regular environmental (ecological and social) challenges" (72).

What is also really important is the fact that the whole book tends to emphasize that the evolution of the mind in humans and nonhuman primates takes place as a mosaic of developmental systems, by seeing populations as ever recurring of stable resources (genetic, phenotypic and environmental, 5) and transforming shared mammalian mental capacities (i.e. *aboutness* as homologous property across mammals). From this latter standpoint, *tEM* is absolutely indebted "to the revolutionary affective neuroscience paradigm" of authors' mentor Jaak Panksepp, especially regarding his conceptualization of the common emotional systems in all mammals who share seven foundational affective systems: fear, lust, care, play, rage, seeking and panic/grief (7-9; 28;33; 37-38; 73). Finally, the top layer of the mindbrain (tertiary level), responsible for

cognitions (language, symbols, executive control and future planning) is energized by the lower-levels emotion: i.e. the *primary-process layer*, largely housed in subcortical areas of the brain and responsible for sensory and homeostatic affects and the *secondary-process layer*, responsible for social emotions, sculpted by learning (associations and mnemonic schemas) and conditioning (largely upper limbic). At this third level "we arrive at uniquely human emotions" (very high and elaborate level of introspective thoughts and imagination). However:

> Here the emotions are still connected to the primary and secondary processes, but they are intertwined in the cognitive powers of neocortex (9)

Accordingly, primary emotions engage deeper and older brain areas and may be activated without the intervention of the cortex and conscious process (28). Emotions in primary and secondary layers are indeed largely unconscious (9).

tEM's approach to mind, its correlative epistemological orientation and ontology, is developed in nine chapters. Given the extent of the covered topics and their speculative richness, in connection with the large amount of experimental data showed up and the almost impossibility to take a vision of synthesis without (inevitably) missing other relevant information, my decision was to take an overall view of the work than to focus on a detailed enumeration of topics.

tEM's affective approach underlines that mind is saturated with feelings (3) and that "*meaning is foundationally a product of embodiment, our relation to the immediate environment, and the emotional cues of social interaction*" (4). By underpinning the fact of not abstract correspondence between sign and referent, Asma and Gabriel announce, more or less unconsciously, the intention to bridge a gap with the phenomenology (Husserl, Heidegger, Merleau-Ponty: 29; 31; 157; 185) and the contemporary Biosemiotics (152: "linguistic brain is not the best model for thinking about how animals or our hominid ancestors engage with meaning"; 157), for which the relational nature of meaning is described as mutual dependence between body and environment.

The most interesting challenge that this book offers is to sketch a speculative (epistemological) and experimental (ontological) context that may explain the way in which the system of emotions can act and constitute "an information-rich niche for human learning" (4), that is to say, how the animal's world, or *umwelt*, is intrinsically emotional (6).

Chapter 1 "Why a new paradigm" (21-42) offers an historical synopsis of the epistemology of the mind (assumptions concerning the nature of mind) by reviewing the two major methods discussed in the early and the late twentieth century: the behaviorism and the cognitive sciences, by recognizing that beyond the value of both practices, they "do not adequately take into account the role of emotions in the mind" (21).

> […] both approaches reveal different levels of mental functioning, but while the former *(behaviorism)* is not flexible enough to explain the adaptability of the mind, the latter (*cognitive sciences*) is neither subtle nor tender enough to explain the heat of consciousness (21)

Among the various paradigms/metaphors of the mind, Asma and Gabriel recall the *associationism* (ideas are copies of sensations) and its main scholars (J. Mill, J.S. Mill, A. Bain). By adopting the idea by which the mind is essentially passive because "it reacts through conditioned reflexes" (23), like association laws, habits, relations of similarity and contiguity, by the work of J. Watson, D.Hebb and B.F. Skinner, this approach developed directly into the *behaviorism*, which focused on stimulus-outcome relationships (23). As argued by Asma and Gabriel, even in the most recent contributions, behaviorism considers emotion secondary with respect to behavior, maintaining the idea that "conditioning is crucial for social cognition and emotional learning" (23).

The emergence of the cognitive paradigm coincides with the advent of the computer era and the rise of a "rational *geist*" (24); the instrumental *ratio*. Behaviors would correspond to internal information states. This is the idea of the computational mind that aims to discredit the cognitive value of other mental processes necessary to the evolution of the human mind: affect, context, culture, history. As *tEM* points out, critics of cognitive sciences underline "its nonbiological approach, its reductionism, and its disinterest in phenomenology and ecological context" (25).

Conversely, post-Darwinian New Synthesis and the Extended Evolutionary Synthesis have elaborated accurate explanations about the adaptive and biological value of the behavior of the perceiver (humans/nonhuman animals): as pointed out by the perceptual psychologist J.J. Gibson, indeed, the environment can be detected "by perceptual systems toward action-responses" (26).

On the other hand, *tEM* also shows how psychology has so far treated the role of emotions ambiguously: emotions can be conscious or unconscious; furthermore there is uncertainty in explaining the generative process leading to the development of properly human emotions and those that would preserve homologous characters with other species (27).

> The major contribution of our book is to put forward a philosophy of affective neuroscience that clarifies the exact role of emotions in a way that may orient future empirical works" (27)

As a case study of emotional intelligence in a prosopagnosic patient seems to suggest (34-37), "perception and affect are bound and actionable before tertiary-level conscious appraisal" (36), by arguing that the affective information in the experiment functions as an unconscious form of recognition (321, note 66).

The experimental results provide evidence that, notwithstanding MJH[1]'s overt non-recognition (i.e. lack of awareness of identity), information about the face's identity is available to an affective reaction system (35).

CHAPTER 2 "Biological aboutness" (43-73) consists in a brief conceptual history of teleology before Darwinian revolution. Especially, Asma and Gabriel focus on three teleology traditions, that is, three types of teleology that are "logically distinct, but the history of biology reveals profound confusion among them" (44): 1) natural teleology as opposed to theology and vitalistic obscurantism, 2) teleology that argues for goals in natural processes and for a naturalistic paradigm of matter's self-organization (i.e. Aristotelian entelechy and autopoiesis) and 3) teleology that argues for searching of goals inside agents (biological aboutness or intentionality).

> This last tradition explores goals that guide animal behavior and can be of two major types: a) representational and b) non-representational. [...] Our claim is that there are at least two forms of non-representational intentionality: (i) perceptual affordances and (ii) affective or emotional intentionality (45)

Perceptual affordances are discussed in chapter three, while affective or emotional intentionality in chapter four.

tEM's attempt is to present a paradigm of teleology in terms of a post-Darwinian reconsidered ontology (14). The major purpose is indeed to capture the truly remarkable feature of the mind, namely its teleological orientation by emphasizing "its embodied active involvement with unique ecological context" (65).

> We need a theory of mind, then, that does not deny intentionality to mind by stipulating purely mechanical or computational modules sculpted by external forces. We also need a theory that does not idolize the mind as a mystical layer of Cartesian consciousness (64)

With reference to this point, as I have already mentioned, I feel a little bit constrained with the simple taking for granted "the mystical layer of Cartesian consciousness", hypothesis that certainly encounters a not insignificant problematization in light of the treatise on *The passions of the soul* (1649).

As the authors suggest, "intentionality is affective firstly – grounded in the adaptive emotions – and only derivatively ideational. [...] As such, aboutness is a homologous property across the mammalian clade, and probably all the way down the chain of biological phylogeny" (66).

The reasons why the authors believe that affect can be (not consciously) intentional are the following:

[1] A patient diagnosed with face-blindness or PA.

1) affects are adaptations (phylogenetically and behaviorally) to environment and challenges. "They are about this problems" (72);
2) affects are mediating and motivating causes, such as lust, seeking ("affects target goals unconsciously when homeostatic imbalances encounter specific environmental conditions", 72);
3) affects have the unique intentionality structure that places their raison d'être outside themselves.

Finally, Asma and Gabriel recognize the existence of some affects that are referential, that is, which have a classic conscious structure, as emotions that agents are aware of (73).

CHAPTER 3 "Social Intelligence from the Ground Up" (74-90) stresses the need to sketch a theory of cognition by arguing for a deeper understanding of the evolutionary processes. Also it suggests a model of social intelligence that relies on perception and affect (75).

> These social interactions are embodied – the perceptual system being the mode in which they occur – and require motivation from the affective systems. Whether they are conscious or not, social behaviors constitute a type of intelligence insofar as they demonstrate integration of knowledge about the past, [...] the present situation and an appropriate understanding of the consequences of action for the future (75)

Accordingly, *tEM* argues for an explanation of both social and emotional intelligence as embodied systems that firstly and promptly require motivation from the affective systems: primary emotions (seeking, rage and fear), secondary/social emotions (lust, care, panic and play) and tertiary cognitive emotions (angst and aesthetic feeling) (77).

Stemming from the fact that social animals need of 1) communicating their homeostatic states, 2) these internal needs are externalized by perceptive and motor equipments (body movements, gestures, sounds, facial expressions) and 3) that they are equipped with an affective system to mediate reception and communication of events, Asma and Gabriel infer that social intelligence may be described as a unit arising from the intertwining of homeostasis, bodily display and affect (82).

CHAPTER 4 "Emotional Flexibility and the Evolution of Bioculture" (91- 121) offers a very interesting comparative analysis of some primate social behaviors – three primary emotional systems such as seeking, lust and care – as resulting forms of bioculture intelligence (118) and embedded in their ecological niches. Especially, seeking, lust and care are described in a context of precognitive notion of social intelligence; that is, as prosocial affordances and affective systems (98). As *tEM* suggests, the seeking system would be classed as a master emotion and truly motivational system. Among the main targets of searching

behaviors, authors make reference to hunting, foraging, procreation, exploration of the environment, the act of paying selective attention (97). Seeking may be outlined as a subjective feeling that "matches those homeostatic imbalances that drive the organism toward resource exploitation and satisfaction" (97). Especially referring to hunting, recent data from "comparative ethology of different primate species and other mammals shows that cooperative hunting does not require cognitive sophistication" (99).

Definitely:

> many behaviors that look cognitively coordinated, like chimpanzee hunting parties, can be explained sufficiently by affective/emotional systems (like seeking), which are channeled by ecological and cultural constraints into dedicated action patterns. Early human seeking is not a different kind of process, but it received its own cultural channeling and evolved into a feedback loop of social learning (99).

In humans, the emergence of such cultural expression of seeking produced the unique effect of (the culture of) curiosity (99; 104). Accordingly, *tEM*'s model requires that the "affective system can be decoupled from their dedicated targets and recruited in new functions, ultimately giving rise to cultural loops" (114). Asma and Gabriel's major objection to cultural evolutionists is that they "have not sufficiently factored emotions into their model of cooperation and group commitment" (116) when really, "affective neurosciences shows that individual mammals already display deep group commitment from the very start, via the care system and imprinting" (117).

Chapter 5 "The Ontogeny of Social Intelligence" (122- 152) aims at describing the ontogeny of social/emotional intelligence that we share with other animals, through the infant-primary caregiver relationship and the developmental impact of early experiences.

In my view, there are at least two very remarkable observations stemming from this analysis: the former concerns the assumption that one of the main important element of our social nature – trust and its cognitive meaning ("probably assembled in the late Pliocene or early Pleistocene period", 123) – , could be considered in terms of "an exaptation of an ancient psychological mechanism" (123). Thus, the overwhelming evidence for the fact that "mimicry and the evolution cooperation provide guidance for rethinking a causal story of what makes humans and human culture unique" (123). Beyond the several neuroanatomical changes that paved the way for the "unique ontogeny of social intelligence in humans", indeed, "homo sapiens emerged with some of this social intelligence already pre-adapted from our ancestors" (124).

The latter consideration is about oxytocin, found only in mammals (110). Especially, the oxytocin system is described as a paradigmatic example of a

plastic and adaptive interface between nature and nurture (130-133), since it "plays an important role in priming mammals to form social bounds, but in turn, the early social environment may also be able to shape the development of the oxytocin system" (132), so much to produce, in adult life, possible dysfunctions in social intelligence as a result of an inappropriate infant-caregiver relations (as demonstrated in both schizophrenia and orbifrontal-damaged patients, 134). Finally,

> Recent understanding of phenotypic plasticity (genetic flexibility in response to environmental change), neuroplasticity [...] and epigenetics (heritable gene-expression switching) have restored developmental biology to a place of pride after z long twentieth-century romance with molecular biology (139).

In CHAPTER 6 "Representation and Imagination" (153-183), Asma and Gabriel describe two main transitions: 1) the first from "perception's automatic behavioral affordances to bodily simulation for action and perception in spatial navigation" (16); 2) the second from "affective reconsolidation of memory in dreams to conceptual and linguistic symbol systems" (16) engaged by voluntary and involuntary imagination.

Definitively, how affect came to be decoupled from its primary and secondary-level functions? And how this led to *sapiens*'s unique cognitive realm of symbols and executive thoughts, to well-structure conceptualization and categorization? To put it another way: "how representational abilities were decoupled from perceptual tasks", thus expanding possibilities for (imaginative) simulation[2] and cognitive abilities and maps? (153; 161-164; 168-183).

> Eventually, representational processes and intentionality evolved atop and interleaved with the affective mind, and the whole nature of the equation became transformed by new dialectical interactions between newer and older parts of the mind (156).

Against nativists, *tEM* emphasizes that Gibson's ecological psychology (Gibson 1977; 1979) seems to offer a very interesting model to understand the relation between mind and niche in bioteleological and bioteleosemantics terms (157). Accordingly (Gibson 1979), affordances[3] and effectivities should be intended as implying one another:

> Affordances are dispositions given by features of the perceived environment to support behaviors, and effectivities are a given animal's dispositions to undertake afford-

[2] Asma and Gabril distinguish between "weak" simulation mode, allocated in Pleistocene, and « strong » simulation faculties partaining to Upper Paleolitich (179).

[3] *Ibidem*, 160: affordances as "imperative forms of informational transfer between creatures and environment", perception and action.

ed behaviors in the appropriate circumstances. Effictivities complement affordances in an informational coupling between perceiver and perceived; [...] proprioception and exteroception imply one another (157).

Since affect may be interpreted as a mode of presentation accompanied by intentionality as based-niche/environment component, the authors are proceeding to describe a possible evolutionary history for the decoupling of affect from its "here-and-now functions" (so called "offline processing of information") to enable (and mediate) complex representational functions (voluntary/involuntary imagination, symbolism, abstraction, referential thought, bonding sense/referent/representamen, etc...). Decoupling is defined "the process that cleaves present-tense perceptual indicative percepts from instrumental proto-beliefs" (159).

One of the pivotal suggestion, supported by the mentioned experimental data is that affect provides a motivating internal context, thus playing an important role in promoting concentration, selective attention and memory retention (163). Against the modular computational model, approaching to mind through essential information-processing, *tEM* recovers and put at the heart of its proposal the fundamental "action" of the mind (172), by considering the elements of body grammar mediated by the cerebellum and the empirical evidence that, albeit "most bodily sequencing may be simple stimulus and responses, it can also be decoupled from immediate stimuli" (174). As decoupled, "sequences must reside in the loop of muscle memory, ecological trigger, and affective intentionality. We might think of these motor sequences as 'premodern concepts' because they are not linguistically grounded, but they have the potential for organizing kinds of experience. Procedural memory, for example, is a form of implicit (often unconscious) memory that consolidates motor responses in long-term memory" (174).

Our ability to coordinate our bodies into sophisticated action sequences, such as in rhythmic entrainment or tool use, stems in large part from cerebellum. [...] Primate cerebella, especially ours, are not just relatively larger than in other mammals but also extremely dense in neuronal connections. [...] Using a comparative study of monkeys and apes, Barton discovered that cerebellum evolution happened six times faster in apes than in other primates. Gorillas, chimpanzees, and humans had a rapid cerebellum expansion that might be uniquely important for explaining our unique mental and cultural advances. [...] The cerebellum is important in modelling, predicting, and organizing behavioural sequences. [...] It is also important in fine-visual-motor dexterity. [...] The ability to string together such behavioral steps is facilitated by cerebellum (not higher cognition) and it makes social learning possible, but is also improved by social learning. We consider the elements of body grammar mediated by the cerebellum to be an important element in the manifestation of the action-oriented representations discussed above (172; 173).

In CHAPTER 7 "Language and concepts" (184-203) is presented the entanglement between emotion, language and concepts, by proposing a biological approach to emotion and an evolutionary comprehension of language in the context of its affective social value (186). Imagination and language co-evolved as much intertwined systems, by assuring higher monitoring and control abilities over internal affectivity (195). As outlined in chapter 6, Asma and Gabriel's suggestion is that "image-based thinking may have dominated our prehistory and formed another domain of premodern concepts, but such a modality is still with us, albeit obscured by the propositional dominance of modern mind" (176).

In CHAPTER 8 "Affect in Cultural Evolution" (204-263), the examination of social structures is considered by the role of affect in evolution, showing the way how societies relate to affective forces. Evolution of society is viewed as based on three stages of social institutions, all present and nested in contemporary society, and suggesting an evolutionary model to interpret the rise of collective behaviors : 1) the basic economic unit of nuclear families, 2) regional groups, as it happened with agrarian model society (due to intensification of economy, technology and sedimentary organization) and 3) the urban global society. "As social institutions become a part of our lived environment" – Asma and Gabriel claim in the introduction – "culture serves as a secondary niche for the species" (17), then arguing the key role of affective adaptation to the specific ecological and cultural niches as causal factor in transformations of social norms .

Reaching the apex of a very speculative pyramidal reflection, in CHAPTER 9 "Religion, Mythology and Art" (264-314), Asma and Gabriel explore the evolutionary paths associated with the emergence and the assembly of transcendental and spiritual emotions and of the all variety and complex range of affective-based adaptations and exaptations assigned to the evolution of art and religion as possible responses to sociocultural problems. "Systems that culturally manage our emotions, like religion, were selected for because they helped early mammals flourish" (19). Behind the feelings of wonder and curiosity, Asma and Gabriel argue about the emotional landscape to explain how moving "from basic affective sources like the seeking and play systems, these spiritual emotions functioned to temper intense feeling of fear and grief in the context of the noecortical imaginative elaboration of culture" (20), then contributing to emotional-based strengthening of the bonding among individuals and groups.

The book is provided with an excellent set of REFERENCES (365-412) and NOTES (317-363), the latter rich in itself of important details that the authors evidently considered not advisable to be introduced in the body of the text, to

avoid weighing it down excessively. Actually, this critical apparatus constitutes a theoretical and speculative appendix complementary to the text, nearly to constituting a continuation of the history, often providing clarifications and historical explanations of the background and suggestions for possibly future research programs.

Elena Pagni
elenapagni@gmail.com
Institute of Human Sciences, Federal University of Juiz de Fora, MG, Brasil

Florence Burgat
Être le bien d'un autre
Rivages, Paris 2018, pp. 128

Gabriele Tassinari

What does it mean "to be the goods of another", "être le bien d'un autre"? The question upon which Florence Burgat constructs his book does not, as it might seem, concern the sphere of moral beneficence or attachments.

The meaning of "bien" examined here is the legal one of property, i.e. an entity that can be used and appropriated for human ends.

Historically, all non-human animals have been subjected to the category and regime of properties. They have been famous victims of the sharp division between persons and *res* operated by the Roman tradition and inherited by most modern legal systems (Korsgaard 2013: 629).

In the first part of the essay Burgat examines this dichotomy from the perspective of the fictitious nature of law and her capacity to establish a universe of discourse independent from natural reality, although designed to order it.

Burgat reflects on the fact that the classification of animals as *res* does not originate from a cartesian belief about their ontological nature. It is clear that animals, from the point of view of natural sciences and phenomenology, cannot be considered as mere objects, deprived of a psychic life, independent needs and a capacity for action. Nevertheless, they can be treated "as if" they were so in the autonomous space of the juridical formulation in order to sanction their economic use.

To underline the independence between the two levels of the speech it suffices to consider the number of different regimes under which an animal of a particular species, for example a rabbit, may fall depending on whether it is classified as wild or domesticate, as pet, farm or research animal.

Legally, biologically similar animals can be classified as quite different objects: the protection afforded to their fundamental interests varies accordingly and to a substantial degree. It could be noted that human beings can also be subjects of different rights depending of contextual and relational ties, such as in the case of citizenship rights. Will Kymlicka has suggested a similar recognition of group-differentiated rights to animal communities (Donaldson and Kymlicka 2011). However, no adult human being, except the

ISSN (print) 2281-8618-ETS

slave, is absolutely excluded from being a subject of justice and from enjoying some fundamental negative rights. The analysis carried out by Florence Burgat is to be understood relatively to this foundational dimension.

The dichotomic classification of persons and properties operated by the juridical system accomplishes therefore a reification of the animal, which is resolved in the utilities and goods that human beings acquire from its body and its activities. This process, however, also influences also the ontological conception: the invisibility of animal subjectivity extends from the juridical fiction both in the realm of empirical science and in that of public discourse.

The second part of the essay aims to support this approach through an in-depth historical examination which I will only briefly mention: on the one hand Burgat reconstructs the evolution of the legal status of the slave in Roman law and in the Code Noir promulgated by Louis XIV in 1685, on the other hand, the evolution of animal law in the French jurisprudence and civil code, from the beginning of the 19th century until today. Several interesting analogies emerge, for example the definition of both slaves and animals as "movable properties", category that includes indifferently inanimate objects that can be moved, and animated ones which are able to move and act on their own initiative. With regard to these objects it exists for the owner the ius utendi and abutendi, which qualify as the right to all that is produced by his property and as the right to sell, dispose or destroy the property itself.

The animal's complete appropriation is realized by its consumption. As for the slave, even if it is not used as food, the right to dispose of it is originally equally unlimited. His life can end with the exhaustion of his productive forces and his killing is punished only to the extent that it constitutes a damage to the owner's belonging.

The evolution of the law sees a gradual limitation of the ius abutendi of the owner towards his "mobile property". The old or sick slave can no longer be abandoned, can be transferred to others in case of excessive mistreatment, the punishments and mutilations that the master may inflict in response to disobedience or attempted flight are defined by the law and the unjustified murder is theoretically prosecuted. Similar limitations of property rights are found in the evolution of animal law, starting with the loi Grammont in 1850, which punishes the mistreatment of domestic animals, although initially limited to acts performed in public.

Two main themes emerge from the analysis of the present situation of animal law: firstly, the increasing recognition of non-human animals as sentient beings endowed with their own interests in the European and French context. Examples are easily found in Article XIII of the "Treaty on the

Functioning of the European Union", in art.515-14 of the French *Code civil* reformed in 2015 and, more implicitly, in the evolution of criminal law. Secondly, the inconsistencies and conceptual contradictions that emerge within the corpus of laws between this principle and their persisting treatment as properties and commodities. The same article of the *Code civil* mentioned above, that states that "les animaux sont des êtres vivants doués de sensibilité", confirms at the same time that "les animaux sont soumis au régime des biens".

The main concern of animal law remains to establish the means, the time, and the purposes for which an animal may be killed or damaged. While the law severely punishes cruelty to animals, understood as perverse infliction of unnecessary suffering, it is reluctant to censure suffering that is functional to the exploitation of animals and integrated into established cultural and industrial practices. The animal, notes Burgat, at the same time belong and does not belong to itself, and his master is by contrast at once owner and guardian.

Starting from these legal and historical assumptions Burgat develops the most philosophically original contributes of the essay. The first one concerns the analysis of the logic that governs the situation of objects reduced to "being the good of others". This logic, as it has been seen, originates within the legal representation but translates in the concreteness of the productive processes of exploitation.

The fundamental feature of this condition is the negation of the personal ends of a subject, understood as the set of his interests and the ability to pursue them – along the lines of what Tom Regan has called "preference autonomy" – and their replacement by the purpose imposed by the owner. Employing Hegel's analysis from the *Elements of the Philosophy of Right,* Burgat shows how the process of appropriation takes place in relation to economic use. His thesis is that the appropriation of the totality of the working time and of production of the subject gives the master ownership over all there is of substantial in him, including his very personality.

Burgat rejects the possibility of a slave's internal, metaphysical freedom, safeguarded independently from all external conditions. In any case, this escape route is much less available for animals as their inferior endowment of higher mental faculties makes them more vulnerable to the immediate reality of their own physical conditions.

Hence, the whole of the capacity of action of a subject is identical to his very being: "prendre toutes les forces d'un individu, c'est le prendre lui-même". The difference between ownership of the use of a subject's abilities and the ownership of the abilities themselves resides therefore only in the limitation of the time and the methods of utilisation.

The thesis is interestingly connected to the debate on the conditions of moral permissibility of at least some forms of animal use. For example, it seems to indicate that the mere protection from suffering and negative mental states is insufficient and should be supplemented with the time and possibility to engage in non-productive activities the animals themselves find rewarding (such as environment exploration, food acquisition, socialization and play). A parallel but descriptive shift is underway in the scientific comprehension of animal welfare, from the narrow hedonism and the traditional five liberties of the Brambell Report toward the appreciation of positive experiences acquired through autonomous activity. Unfortunately, the analysis of Burgat does not fully clarify the underlying normative reasoning, which seems reasonable however to link to a deontological approach, which could match well with Martha Nussbaum's perspective on dignity and capabilities.

Either way, the book is not limited to the *pars destruens*, but leads to a normative position. In general the solutions regarding the inconsistencies of the juridical status of non-human animals can be divided into three main groups: a pragmatic approach suggests to keep the non-human animals in the legal category of goods, although as special objects of specific protection duties defined through the resources of ordinary legislation (Favre 2010). A second approach, adopted for example by Christine Korsgaard, involves the redefinition of animals as subjects of law, through the creation of one or several intermediate categories between persons and properties. Lastly, a third and more radical proposal advocates the recognition of their subjectivity through full inclusion in the category of persons.

The author endorses the last option. Two main arguments are offered: firstly, the autonomy of the juridical formulation is recalled in relation to the vexed question of the qualities or dispositions necessary to be considered a person. Burgat notes that kantian or contractualist theories of personality, more demanding in terms of rational and linguistic requirements, involve the exclusion of many humans from the set of persons. From this observation the well-known marginal cases argument arises. However, Burgat is here rather interested in the process whereby marginal cases are protected "as if" they were persons even though they ontologically are not, specularly to how the slave could ontologically but not legally qualify as a person. The conclusion is that legal personality is not a quality that can be verified, but rather a prescriptive stipulation: "le fruit d'une décision à visée morale, protectrice, dispensatrice de droits, alors qu'aucun accord n'existe sur la compréhension et sur l'extension du concept". Hence, the law may include or exclude objects from the category "à partir du moment où la mentalité d'un société l'y invite fortement".

The second argument concerns the moral importance of sentience and of being the subject of an individual psychic life for having the right to the protection of one's fundamental interests. The will to protect these conscious lives, as already expressed by the legislator, along with the dichotomic division of the juridical system and the plasticity of the legal instrument, contribute to the extension of personality to non-human animals.

The topic is well elaborated throughout the essay and relies on examples of legal instances aimed at the recognition of the *habeas corpus* mainly in favour of great apes, some of which, such as the case of chimpanzee Cecilia in Argentina in 2016, were successful.

However, at least two questions remain to be answered.

On the one hand, how the substantive question of the moral status and standing of animals is to be resolved in the context of a society's political culture, to which the legal instruments refers. Burgat's argument about the plasticity of positive law establishes the possibility, not the necessity, of the inclusion of animals in legal personhood. A change in public culture and mentality is also required, but this change is far from realized. It is important to remark that the recognition of animal sentience in law and culture can have different meanings according to the underlying theories of duties towards them. At present, the most shared conception is probably that humans have, at best, a duty not to inflict pain and suffering to animals but not a general duty not to kill. Robert Garner, for example, believes that for now a right to life should be confined in the domain of ideal political theory, in favour of a non-ideal position focused of the rigid prohibition about the infliction of suffering (Garner 2013).

Secondly, besides public consensus, there are genuine normative questions about the precise set of fundamental rights which legal personality should protect in the case of animals. Fundamental rights are effectively the same, not only for men and other beings, but across all sentient animals? It has been argued that most sentient animals have a right not to be harmed or killed but do not have a genuine right to be free and not to be owned (Cochrane 2012). This position would not be fatal to Burgat's argument – after all, personhood is awarded also to non-autonomous humans – but surely requires to be explicitly engaged.

Gabriele Tassinari
gtassina@unito.it
University of Turin - FINO Phd Program
École normale supérieure of Paris

References

Cochrane, Alasdair, 2012, *Animal rights without liberation: applied ethics and human obligations,* Columbia University Press, New York.

Donaldson, Sue and Will Kymlicka. 2011, *Zoopolis a Political Theory of Animal Rights*, Oxford University Press, Oxford.

Favre, David, 2010, "*Living Property: A New Status for Animals Within the Legal System"*, in *Marquette Law Review*, 93: 1021-1069.

Garner, Robert, 2013, *A theory of justice for animals: animal rights in a nonideal world*, Oxford University Press, New York.

Korsgaard, Christine. 2013, "*Kantian ethics, animals, and the Law"*, in *Oxford Journal of Legal Studies*, 33(4): 629-648.