The University Of Sheffield.

BARBARA VETTER'S POTENTIALITY ACCOUNT OF COUNTERFACTUALS

Giulia Casini

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

The University of Sheffield
Faculty of Arts and Humanities
Department of Philosophy

5th October 2022

ACKNOWLEDGMENTS

I wish to thank my supervisors Dominic Gregory and Rosanna Keefe for their support and help over the years. It was a long and sometimes difficult journey, but it would have been even more difficult without your kindness and encouragement. I will treasure your teaching for ever.

I want to thank the whole Department of Philosophy at the University of Sheffield, both those who are still there and those who left, for creating such a great community and safe space for everybody to work. The insight and friendship of other students and the dedication of all staff were invaluable in the good times but even more in the bad ones.

I must thank *Argumenta*, the Journal of the Italian Society for Analytic Philosophy, and the editors of its Special Issue on *The Source of Modality*, Giacomo Giannini and Joaquim Giannotti, for giving me the opportunity to publish some of the present work in that issue. My gratitude goes also to the referees, whose comments on my article about would-counterfactuals inspired some essential discussions which greatly improved the quality of this thesis.

Thank you, Nick, for believing in me, even when I didn't, and for showing me the light when everything seemed dark.

Thanks to my parents for their trust, and to all my family both in Italy and England for their love and warmth.

To Renato and Mario,

who inspire me every day,

in their own different way.

ABSTRACT

This study comprises an analysis and an evaluation of Barbara Vetter's proposal for an account of counterfactuals in terms of potentiality. Counterfactuals are a kind of conditional, but they are now largely recognised as part of modality and as such it is thought that an account of them can be developed from a theory of possibility. The traditional metaphysical background of counterfactuals appeals to a theory of possibility based on possible worlds. In this study, we wish to explore a different theory in the hope of finding a different, but equally effective, background. This theory of possibility is Vetter's potentiality account of modality. The appeal of such an account comes from the idea that potentiality is a class of properties of the objects of the actual world, so that an account of modality based on potentiality could intrinsically link modality to the actual world. If counterfactuals can be successfully inserted in this account, then they too can be concretely linked to the actual world and avoid the metaphysical problems connected with possible worlds. However, the extension of Vetter's framework to counterfactuals proves challenging. The scarcity of details on the matter in the original material makes the investigation more difficult and certain pressing questions about the applicability of the whole framework arise from it. The resulting critical discussion of potentiality and counterfactuals is hoped to indicate potential ways to further elaborating Vetter's picture, rather than grounds for rejecting it, in such a way that the view can accommodate counterfactuals more effectively and, by doing so, provide a more complete account of modality.

LIST OF CONTENTS

DECLARATION

I, the author, confirm that the Thesis is my own work. I am aware of the University's Guidance on the Use of Unfair Means (www.sheffield.ac.uk/ssid/unfair-means). This work has not previously been presented for an award at this, or any other, university.

Giulia Casini

INTRODUCTION

The aim of this thesis is to discuss and evaluate Barbara Vetter's proposal for an account of counterfactuals as presented in her book *Potentiality: From Dispositions to Modality* (2015).

Counterfactuals play an important role in our everyday lives: our ability hypothesise about how things could turn out be, or about how things could have been different, is essential both for planning our actions and for learning from our past mistakes and successes. Counterfactuals are also at the core of the scientific investigation since scientists come to design their scientific experiments by thinking counterfactually.

The importance of studying and analysing counterfactuals has been recognised by philosophy relatively recently, from the 1940s and 50s onwards, and the pursuit of such a study presented several challenges, as it still does. The discussion of counterfactuals goes well beyond semantics, spacing from epistemology to ethics, from logic to metaphysics.

Several theories of counterfactuals were developed over the years, but there is no doubt that one of the most popular is the possible world account of counterfactuals, developed first by Robert Stalnaker and then by David Lewis. This account has indeed an unquestionable appeal, for its intuitive simplicity (the idea of truth based on similarity with reality is in many ways a stroke of genius, at least at an intuitive level) and for the naturalness with which it can be inserted in an already well-established apparatus (by being an extension of the logic and the semantics developed for possible worlds).

Nevertheless, such an account throws counterfactuals into the ontological storm surrounding possible worlds. Even if the account of counterfactuals itself is enticing and quite compelling, what encircles it is almost like a fishing net that entraps it in a much broader and

almost unsolvable debate. The possible world account of counterfactuals cannot escape from the question "What are possible worlds?", so that what at the start seemed quite an effective semantic picture, turns into what is for some a metaphysical nightmare. Possible worlds are wonderful tools until you try to understand what they are.

In truth, counterfactuals cannot avoid the metaphysical debate: they are, after all, about how things could be different, and any account involving an investigation of how things could be different from what they are needs to answer certain metaphysical questions. In this sense, we agree that counterfactuals are part of modal discourse: suggesting ways in which things could be different is simply a matter of conceiving the *possibility* for these things to be different. And the discourse on such a possibility ends up necessarily in a metaphysical debate on the nature of possibility itself. However, maybe the kind of metaphysical debate that possibility and counterfactuals encompass does not need to be about possible worlds. Maybe we can find other ways to explain how things can be different, rather than say that they can be so because they are so in other worlds.

Here is where Barbara Vetter's work comes into the picture. Her book devotes itself to develop a new metaphysical background for possibility, and her plan is to do so not by looking for answers in other worlds, but by grounding possibility in the objects of our own actual world. Potentiality is at the core of this new background, by being at the same time the source of possibility and a class of real properties of actual objects. This is the appeal of her theory: the idea that we do not need anything other than the objects of our world and their properties to explain possibility, and to explain counterfactuals too.

To get such a result would be splendid for all those who are sceptics about possible worlds, and have a certain concrete idea of reality, but do not want to give up modal discourse. However, Vetter's proposals about counterfactuals in particular amount only to a small part of her project overall.

Nevertheless, the potential of Vetter's ideas, and the appeal of her metaphysical picture overall, makes it well worth seeking to develop an account of counterfactuals based on potentiality. Our investigation will try to explore how the sketch that Vetter offers could turn into a more complete account of counterfactuals. Unfortunately, we will discover that there are some obstacles and difficulties standing in the way of this approach. To overcome these problems, a much longer work would be needed than what I can offer in this thesis. But to point

to the issues is already a step forward towards the identification of possible solutions. Once a problem is known, it gets much easier to tackle.

My hope is that the ideas presented across this work will not lead to a rejection of Vetter's account but rather to the articulation of ways of developing it so to include counterfactuals in a more extensive and effective way. I suspect that Vetter's project is valid not only as a theory of possibility, but also as a theory of counterfactuals. For this to be proved though, there is much that still needs to be done.

This thesis will have the following structure.

The first chapter will be an introduction to counterfactuals: we will discuss the relation between counterfactuals and other conditionals, and we will introduce some of the types of counterfactuals that we will discuss across this work.

The second chapter will clarify the connection between counterfactuals and modality, by explaining roughly what modality is, by introducing certain ways to categorise it, and by showing some of the problems that accompany modal discourse.

The third chapter will move towards Barbara Vetter's account by presenting dispositionalism in general, by suggesting some of the reasons why we prefer it to a possible world account of modality, by explaining how Vetter's potentialism is different from the traditional versions of dispositionalism and finally by presenting her proposal for an account of counterfactuals.

The fourth chapter will focus on those details of Vetter's account that are essential to understand her proposal concerning counterfactuals, both in terms of the metaphysics she develops based on potentiality and in terms of the semantics she attaches to it.

The fifth chapter will finally present my own proposals on how different kinds of counterfactuals could be accounted for by expanding Vetter's initial suggestion, so that we will see what is needed for these proposals to work.

The sixth chapter will conclude the thesis by presenting the problems that arise from Vetter's account of counterfactuals, and the associated notion of 'being an earlier stage', and by showing how these problems could be potentially damaging for her whole picture.

The conclusion of this work will point out these risks, but it will also try to end on a positive note, suggesting that, in the light of the findings offered by this thesis, Vetter's project

is still well worth pursuing, but in a way that keeps a closer look to the role of counterfactuals and that clarifies several points concerning iterated potentiality.

CHAPTER I

COUNTERFACTUALS

The aim of this thesis is to discuss an account of counterfactuals based on the potentiality framework developed by Barbara Vetter. An essential starting point for pursuing this task is to define what counterfactuals are. This will be the objective of the present chapter.

Counterfactual reasoning plays a huge and crucial role in our mental lives. In planning what to do and explaining what has occurred we think and say things like "If A had not happened, then B would have happened". Therefore, even if we might not realise it, counterfactuals are a constant presence in our everyday life.

Counterfactuals are conditionals. However, they seem very different from other sentences that we also call "conditionals". In this chapter, we will try to make sense of this difference and explain why we cannot apply to counterfactuals the same rules that we would use for other conditionals.

Although most counterfactuals share the same features, we can identify various types of them. Different kinds of counterfactuals may require different approaches or be problematic for an account aiming to explain counterfactuals exhaustively. In this chapter then, we will address some of these "special" types of counterfactuals, showing in which way they diverge from the others, and which challenges they might bring about.

The structure of the chapter will be the following. In the first section, we will introduce conditionals in general, focusing on the conditional structure and comparing it with the counterfactual structure. We will also introduce the relation between antecedent and consequent, which will be a recurring theme of this work. The second section will list three

main kinds of conditionals: material, indicative, and subjunctive, and will discuss in which way subjunctive conditionals (that is, counterfactuals) are different from the other two. There will be a short digression on the distinction between the names 'subjunctive conditional' and 'counterfactual' in the third section, explaining the issue of the falsity of the antecedent. In the fourth section, we will concentrate on the relation between counterfactuals and possibilities, as the need for an account of metaphysical modality is one of the main characteristics of this kind of conditionals, which leads them to be inserted in the potentiality framework. The fifth section will see a taxonomy of different types of counterfactuals, while the sixth will focus on the relation between two of these types, would- and might/could-counterfactuals. The final section will conclude the chapter in preparation for the next one.

1. Introduction to Conditionals

In answering the question on what counterfactuals are, the first thing we notice is that they are conditionals. If we take a simple sentence to be something like "The cat is on the sofa", then we can describe conditionals as complex sentences, meaning that they are the result of a combination of other sentences. In the case of conditionals, we normally see two sentences, one introduced by the particle 'if', and the other introduced by 'then', like "If the cat is on the sofa, then the dog is on the carpet". These two sentences so combined are called 'antecedent' (the clause after the 'if') and 'consequent' (the clause after the 'then'). The 'then' might not always be present, and the consequent may sometimes precede the antecedent, but the standard form of a conditional is the one described. In considering conditionals, we usually do not focus on these minor layout variations, as any conditional can be reformulated in its "If …, then …" form (for an introduction to conditionals see Bennett (2003), Edgington (1995), Sanford (1989), Harper *et al.* (1981), Wood (2003)).

However, there are different kinds of conditionals, which present variations of this form. The discussion of these kinds will be matter of the next section, but for the moment it can be helpful pointing out a difference in structure between counterfactuals and other conditionals. While most conditionals (including what we will call 'indicative conditionals') have the shape: "If A, then C" – with A being the antecedent, and C being the consequent – counterfactuals are considered to have a different structure, even if they still include the presence of "If…, then…". Instead of having the simple structure "If A, then C", counterfactual conditionals have the following:

"If *it were the case that* A, then *it would be the case that* C"

A and C are still two sentences called 'antecedent' and 'consequent' that can have variations in their order, but again, these variations can be ignored. Note that this definition of the counterfactual structure is not exclusive: we consider counterfactual conditionals also those presenting this structure in its past form: "If *it had been the case that* A, then *it would have been the case that* C", and those presenting it in its *could* and *might* variants, including again their past form: "If *it were the case that* A, then *it could/might be the case that* C", "If *it had been the case that* A, then *it could/might have been the case that* C". All these structures are presented here in a relatively lengthy form, used to capture the content that is common to a wide variety of cases that are phrased in superficially different ways. Therefore, in what follows we will tend to use more "common" counterfactuals, like "If John were going shopping, he would get some milk" rather than "If it were the case that John is going shopping, then it would be the case that he gets some milk".

After describing what the counterfactual structure is, it is important to ask which sentences we should take to be the components of this structure. We can find several examples of what sentences A and C could be, both in the simple "If A, then C" structure and in the more complex counterfactual structure. Taking the simple conditional structure for example, we can have conditionals like "If you press that button, the bomb will explode", "If you do not give me back my money, then I will break your legs", "If you have won the lottery, then can you buy me a car?", "If you do not pass the exam this time, then work harder next time!". Even if many different sentences can be placed as antecedent and consequent of a conditional, for the current discussion there are some that we will exclude.

In this work we will ignore conditionals and counterfactuals involving non-assertoric sentences such as imperatives ("Work harder next time!") and questions ("Can you buy me a new car?). Therefore, in our analysis, we will not discuss counterfactuals including questions like "If John went shopping, would he remember to get some milk?", but only counterfactuals that are assertoric, that is the ones in which both antecedent and consequent are assertions.

The attention of the literature has always been predominately on assertoric conditionals and counterfactuals anyway because they have the greatest relevance to our lives: we use them to predict danger ("If I were to touch this baking tray that has just come out of the oven, I would burn my hand"), to explain natural laws ("If water is brought to the temperature of 100 C°, it will boil"), to design scientific experiments ("If I immerse this rock in this glass full of water and

the water spills out, then we can prove that water has a volume"), and for many other purposes. To be able to attribute a truth value to these conditionals and counterfactuals helps us understanding our reality and our world. Therefore, it is extremely important to find out what their truth conditions are.

However, counterfactuals and other conditionals may not have the same truth conditions. To test this, we should compare counterfactuals to other kinds of conditionals and see if the truth conditions usually associated to the latter can work also for the former. This is the aim of the next section, in which we compare three types of conditionals.

1.1. Antecedent and Consequent

Before we move on to the next section though, there is a crucial aspect of all conditionals, including counterfactuals, that we need to discuss. A conditional, like a counterfactual, is very clearly *connecting* the two other sentences that are called 'antecedent' and 'consequent'. However, it is quite controversial what the nature of this connection should be. The name 'conditional' intuitively suggests that the antecedent is some sort of "condition" for the consequent to happen, which, by name again, seems to be a "consequence" of the antecedent, but this intuitive picture does not seem to cover all our uses of the conditional and counterfactual structures.

Defining this connection is probably the biggest issue surrounding counterfactuals, and conditionals in general as well. However, our focus does not need to be on this issue concerning all conditionals, because here we are only interested in counterfactuals. Also, since the focus of this thesis is Barbara Vetter's potentiality account of counterfactuals, we will only need to see how this connection is dealt with in this account – and we will see that the solution of this issue is actually a source of problems for Vetter.

Nevertheless, a general and intuitive view of the connection between antecedent and consequent of a counterfactual can be useful here, also if we consider the distinction we will make later between might/could- and would-counterfactuals. Very roughly, we can say that a counterfactual is used to express the fact that, given that a certain antecedent is the case, then a certain consequent is the *outcome* of this antecedent (whether the only outcome or one out of many outcomes we will discuss later). We can suspend judgment on what we mean by 'being the outcome of', because we only need an intuitive idea of this connection at this stage.

Therefore, we can say that a counterfactual expresses a kind of *dependence* of the consequent on the antecedent, attributing to 'dependence' its most general and "neutral" meaning.

Naturally, this is a very limited view of the relation between antecedent and consequent, which does not really cover all the possible cases of counterfactuals. In truth, to account for *all* counterfactuals is a difficult issue. In particular, trivial counterfactuals ("If I were Italian, I would be Italian) or counterfactuals where the consequent does not seem to depend on the antecedent ("If I were able to play that piano, then I would have at least three fingers") represent a problem and an exception to our intuitive view above.

However, in this work we are mostly interested in finding a metaphysical background for "regular" counterfactuals, rather than discussing the most complex and tricky exceptions. Therefore, we will limit our research to those counterfactuals clearly involving some kind of "dependence" between antecedent and consequent, so that the consequent can be considered an "outcome" of the antecedent, intending both the terms 'dependence' and 'outcome' in their most general meaning.

The importance of the connection between antecedent and consequent will be a recurring theme of this work and will become clear the more we proceed towards a potentiality account of counterfactuals. For now, though, we will focus on the definition of counterfactuals, considering them in comparison with other conditionals.

2. Three Types of Conditionals

Three types of conditionals have often been identified: material conditionals, indicative conditionals, and subjunctive conditionals. Subjunctive conditionals are what we have called 'counterfactuals' previously, so they will be the focus of our discussion. In what follows, I will use the terms 'subjunctive conditional' and 'counterfactual' interchangeably, leaving the explanation of these two names for the next section. To understand what subjunctive conditionals (*alias* counterfactuals) are, as well as what their truth conditions could be, it can be useful to compare them with the other two kinds of conditionals and see in which ways they are different.

2.1. Material Conditional

The first type of conditional to compare with subjunctive conditionals is the material conditional. The name 'material conditional' (or 'material implication') is generally associated with the arrow '→' or hook '⊃' symbol of classical propositional logic and its truth table. The truth table is what illustrates the truth conditions of a material conditional based on the truth values of its antecedent and consequent. A material conditional "A ⊃ B" has the following truth conditions:

A	B	A ⊃ B
T	T	T
T	F	F
F	T	T
F	F	T

We can see that the only case in which a material conditional is false is when A is true, and B is false. Therefore, a material conditional is always true when the antecedent is false. This shows a first, strong distinction between material conditionals and subjunctive conditionals. In fact, while material conditionals are always true when the antecedent is false, we would not think the same of subjunctive conditionals. For instance, we would judge false a counterfactual like "If Lisa had not got a cat, she would have got a dinosaur", even if we know that Lisa has in fact a cat and consequently that the antecedent is false.

The falsity of the antecedent is not a guarantee of the truth of a counterfactual, as it happens for "A ⊃ B". Therefore, we can see that the rules and truth conditions of the material conditional clearly do not apply to counterfactuals. We can conclude that the truth table of the material conditional does not help our search for truth conditions of subjunctive conditionals.

2.2. Indicative and Subjunctive Conditional

If a comparison with the material conditional does not offer clues for the truth conditions of counterfactuals, then we can try to compare them with the second kind of conditional we can identify: the indicative conditional. While the material conditional is found in propositional logic, indicative conditionals are part of the natural language, like subjunctive conditionals, so they might be more useful for our purposes. I will not say anything on the relationship between

material conditional and indicative conditional, as it is beyond the scope of this work, whose main interest is only on counterfactuals.

The names 'indicative conditional' and 'subjunctive conditional' have grammatical origins: they were expressing the idea that, while in the subjunctive conditional the verbs of the conditionals are in the "subjunctive mood", in the indicative conditional, instead, the verbs are in the "indicative mood". These terms are likely to be appropriations from other Grammars, like the Latin or the French, in which this distinction is much more evident. However, if the mood of a verb can be easily identified in other languages, in English these moods are almost non-existent, so this distinction loses most of its original significance.

Nowadays, forgetting the grammatical reasons, the distinction between indicative and subjunctive conditionals is better explained through examples. Therefore, we can appeal to a reinterpretation of the well-known example given by Ernest Adams (1970, p. 90), often used to give account of the difference between indicative and subjunctive conditionals[1]:

(1) If Oswald did not kill Kennedy, someone else did

(2) If Oswald had not killed Kennedy, someone else would have

Case (1) is an example of indicative conditional, while case (2) is a subjunctive conditional. In Adams' intention, we should immediately see that (1) and (2) have different meaning and require distinct truth conditions, so that an analysis of indicative conditionals is not applicable to subjunctive conditionals. In what follows, I try to explain this argument starting from our intuitions on both examples.

In (1) we accept the fact that Kennedy is dead, and someone killed him, so that if it was not Oswald who killed him, someone else must have. Note that even if we are reflecting on the possibility that Oswald might not be the murderer, we do not deny the fact of our world that

[1] Adams' original examples are the following:

"If Oswald had not shot Kennedy in Dallas, then no one else would have"

"If Oswald did not shoot Kennedy in Dallas, then no one else did" (1970, p. 90)

Adams aims to show that in this case the indicative correspondent of the subjunctive is unjustified while the subjunctive is itself justified. Like others (e.g., Lewis, 1973a, p. 3), I prefer to adopt a different version of the example to make clearer that indicative and subjunctive are different even when one of them is not unjustified or absurd.

Kennedy was killed, because we know he was. We may doubt that Oswald is the culprit, but we cannot deny that someone killed Kennedy. Therefore, we are inclined to accept (1) as true.

In (2), instead, we contemplate the possibility that the killing could have not happened altogether. In this case, the focus is not on whether Oswald is the one who killed Kennedy, rather we are questioning a well-known fact of our world, considering the possibility that the killing itself might not have happened. We are less inclined to accept (2) as definitely true, as we would do for (1), because if Oswald had not killed Kennedy, maybe nobody would have killed him, and he would still be alive today or would have peacefully died of old age.

This should show immediately that the two conditionals have different truth values, and so that they must have different truth conditions.

We could say that the reason behind this difference is that, in comparison to indicative conditionals, it seems that sometimes, for judging subjunctive conditionals, the facts of the actual world are not sufficient for determining their truth.

While this might not be evident in (2), there are examples of subjunctive conditionals where we can see more clearly that the actual world does not seem to cover all that we have to consider for determining their truth. Consider the following counterfactuals:

(3) If pigs flew, I would ride one to go to work

(4) If Rome were in Turkey, all Romans would speak Turkish

In this case, we can clearly see that the corresponding indicative conditional versions of these examples cannot work: "If pigs flew, I rode one to go to work" and "If Rome was in Turkey, all Romans spoke Turkish" might not be entirely nonsensical but are clearly wrong.[2] Also, in

[2] Note that one issue here is that, as these corresponding indicative conditionals have false antecedents, those philosophers who hold that indicative conditionals have the same truth conditions of material conditionals would hold that the relevant indicatives are actually true, since a material conditional is true when its antecedent is false. So, the main difference with subjunctive conditionals is once again that, for example, we can have cases in which the antecedent is false, and the consequent is also false, but the subjunctive conditional is true. The actual truth values of the antecedent and the consequent thus do not determine a truth value for the subjunctive conditional, while some philosophers think that it determines a truth value for the indicative conditional. However, we have already said that we will not discuss the relation between material conditionals and indicative conditionals here, so we will not consider this correspondence any further.

both (3) and (4), it looks like we cannot appeal to the facts of the actual world to find out whether these counterfactuals are true or false. To see this, contrast them with the following counterfactual:

(5) If this ginger cat had kittens, they would be ginger too

Even if, also in the case of (5), its indicative conditional version "If this ginger cat had kittens, they were ginger too" seems just as wrong, still we can see that (5) is clearly true in virtue of some facts concerning the actual world, like actual genetic facts about the cat. On the contrary, in (3) and (4) we do not seem to be able to make a similar move concerning their truth. In both cases, we cannot appeal to the same kind of actual facts as we did for (5), because both the pigs' ability to fly and the different geographical location of Rome are not facts of the actual world. As we will explain, this is because counterfactuals like (3) and (4) involve metaphysical possibilities that exceed physical possibility. Therefore, it seems that subjunctive conditionals are different from indicative conditionals because some of them appeal to a kind of possibility that requires a discussion of metaphysical modality (which is intended as the category including metaphysical possibility and necessity) that goes beyond physical modality.

Note that, with this, I am not excluding that there is a relation between indicative conditionals and possibility. My aim here is not to deny any connection between indicative conditionals and possibility in general, or metaphysical possibility, but I am not interested in discussing whether this connection really exists either.[3] What I want to point out is that subjunctive conditionals seem to require an account of metaphysical possibility to go with them because they clearly involve this sort of possibility sometimes, while indicative conditionals do not seem so clearly to require such an account.

Note also that some philosophers have advocated a unified account of conditionals, offering truth conditions valid both for indicative conditionals and subjunctive conditionals.[4]

---

[3] An example of someone supporting the idea that indicative conditionals are also related to possibility is Robert Stalnaker (e.g., 1968).

[4] Including Adams (1975), Dudman (1988, 1994), Edgington (2003, 2008) and Bennett (1988, then revised in 1995). Also, Stalnaker (1968) supports a similar view, but in an alternative way: instead of suggesting an account of indicative valid also for subjunctive conditionals, he thinks that his account of subjunctive should be extended to indicative conditionals.

While I strongly support the idea that subjunctive and indicative conditionals are different and cannot be reduced one to the other, I can see why a unified account of conditionals is desirable. However, the position adopted in this work is that the current accounts of indicative conditionals do not apply to subjunctive conditionals in a satisfactory way.

Even if the potentiality account we will discuss in this thesis is focusing on counterfactuals only, I do not exclude that it could be helpful also in the comprehension of indicative conditionals. I will not judge whether this application is possible here, but maybe the key to the puzzle is starting from a good account of subjunctive conditionals and see if this is applicable to indicative conditionals, rather than vice versa.

In this work, we will not adventure any further on whether a unified account of conditionals is possible, but we will only focus on what could be a good account of counterfactuals. With this section, I wanted to show that the best approach to counterfactuals is to consider them independently of the other kinds of conditionals and that, in producing an account, the best thing is starting from scratch, focusing on counterfactuals only.

3. Subjunctive Conditionals vs Counterfactuals

In the previous section, I kept using 'subjunctive conditionals' and 'counterfactuals' as synonymous. Before moving forward in our discussion, I would like to say few things on these two terms, including why someone prefers using one rather than the other.

The origin of term 'counterfactuals' comes from the name 'contrary-to-fact conditionals' derived from the Latin expression *contra facta*, meaning the 'against the facts'. This name captured a strong intuition we seem to have about counterfactuals: that they are, indeed, contrary to facts, meaning that we are driven to interpret their antecedent as false.

This intuition arises because we very often use counterfactuals to talk about things that we know to be false. Take the following example:

(6) If Napoleon had not been defeated in Waterloo, he would not have been exiled in Saint Helena

In this case, we know that the antecedent is false, because we know that Napoleon was in fact defeated in Waterloo, and so what said in the antecedent is against the facts. It is very

common to use counterfactuals for discussing how things could have been different from how they are, and in doing so we are simply denying how things are, hence going *against the facts*.

In the case of (6), going against the facts means that the antecedent is known to be false. However, there are also many counterfactuals in which we do not know for sure that the antecedent is false, like:

(7) If it were raining outside, I would use my umbrella

(8) If the doorbell had rung, the cat would have hidden under the bed

Even in these cases we may seem to maintain the same intuition that the conditionals presuppose the falsity of the antecedent. We get the impression that the speaker in (7) believes that it is not raining outside and that the speaker in (8) believes that the doorbell did not ring. However, we need to investigate what the nature of this intuition is.

Remember that the aim of this thesis is to work out what the truth conditions of counterfactuals could be. Therefore, we need to enquire whether the intuition that counterfactuals presuppose the falsity of the antecedent is also part of their truth conditions.

To answer we could try to find a counterexample, which would be a counterfactual that clearly does not presuppose the antecedent is false[5]. Such a counterexample would disprove that our initial intuition is an intrinsic semantic feature of counterfactuals because it would prove that there can be counterfactuals raising a different intuition but that we would still identify as counterfactuals. It was Anderson (1951) who formulated the required counterfactual:

(9) If Jones had taken arsenic, then he would have shown exactly the same symptoms which he does in fact show (Anderson, 1951, p. 37)

Examples like (9) clearly show that the intuition of the falsity of the antecedent is not part of the semantic description of counterfactuals and there is no semantic rule compelling us to include this intuition in their truth conditions. We cannot say that (9) is not a counterfactual, even if clearly it is not against the facts.

[5] Note that this does not mean that this counterexample must presuppose that the antecedent is true.

How can we account for our intuition then? Karttunen and Peters (1978) offer a good solution: even if the presupposition that the antecedent is false is not part of the semantic meaning of counterfactuals, it can still be seen as a default interpretation – e. g., a conventional implicature – in the realm of the pragmatics of language. I will not explore this any further, as it would bring us too far away from the themes of this work. However, I think it is important to show that our initial interpretation of counterfactuals as contrary-to-fact conditionals does not need to be rejected altogether, but it simply belongs to another aspect of counterfactuals that is not matter of the present discussion. Nevertheless, for what concerns the semantic interpretation of counterfactuals, I agree to say that they are not contrary to facts.

This is the reason why the use of the term 'subjunctive conditional' has been preferred to 'counterfactual'. Because of its linguistic origin and meaning, it is thought that using the term 'counterfactual' for this sort of conditionals could lead to believe that these conditionals do semantically imply that they are contrary to facts. Therefore, the use of the term 'subjunctive conditional' is usually meant to avoid this confusion, because it does not have any reference to an against-the-facts intuition.

Sometimes in the literature (see Sanford, 1989, p. 77) it can be found that these two terms are used to describe two different things: 'subjunctive conditional' is used to indicate those conditionals that are not indicative conditionals that can have both true and false antecedents, while 'counterfactual' is used to mark those subjunctive conditionals that have false antecedent and consequent, maintaining in this way its original meaning of 'contrary to fact'.

However, in this work, I will keep using the two terms as synonymous. I support the idea that the intuitive falsity of the antecedent is not part of the truth conditions of subjunctive conditionals, but I will still use the name 'counterfactual' for referring to them, also when they have a true antecedent. This is only for simplicity, without implying any reference to the meaning of *contra facta*.

4. Counterfactuals and Possibilities

In this section, we aim to discuss more clearly the relation between counterfactuals and possibility, as this is essential to understand why counterfactuals could be included in modal

discourse and, therefore, in the potentiality framework. Here we will focus on two kinds of possibilities that can appear in counterfactuals.

As said in section 2, counterfactuals sometimes involve 'metaphysical' possibility and because of this they require an account of metaphysical modality to go with them. It is the fact that they require such an account that makes the potentiality framework an appealing background to an account of counterfactuals.

Also, we have seen that there are counterfactuals in their past form, which means that counterfactuals both involve possibilities directed towards the future and possibilities set in the past, so that counterfactuals also involve possibilities in time.

Note that counterfactuals can also involve impossibilities, like "If 2+2 were 5, then (2+2+2) would be 7". The case of counterfactuals with impossible antecedents is quite tricky for any account of counterfactuals, and it will not be discussed in detail in this work.[6] However, we need to make clear that the fact that they include impossibilities does not imply that these counterfactuals cannot be evaluated. Recalling our rough interpretation above, we intuitively evaluate counterfactuals by assuming that the antecedent is the case and then seeing if the consequent is an outcome of this antecedent. Therefore, when a counterfactual has an impossible antecedent, we are allowing this impossibility to be suspended for the sake of the evaluation of the counterfactual and we are taking such an antecedent to be the case, at least at an intuitive level, even if it is impossible.

4.1. Metaphysical Possibility

Defining metaphysical possibility is not an easy task. We can find some very different definitions and philosophers seem to use the term to mean different things. Here I do not want to attempt to define this notion precisely. The most important thing that I want to stress with this section is that counterfactuals are sensitive to possibilities which seem to be not so easily

[6] Counterfactuals involving impossibilities are called 'counterpossibles'. Even if it would be interesting discussing counterpossibles in this work, and in particular some difficulties that they can create for a dispositional account of counterfactuals, as pointed out by Barbara Vetter (2016), we will not do so. As said, this thesis pursues an account limited to "regular" counterfactuals without focusing too much on the exceptions.

explained by appealing to the actual world alone, i.e., metaphysical possibilities, and that we need to find a way to explain these possibilities when we look for their truth conditions, hence the need for an account of metaphysical modality.

The easiest and most intuitive way to define metaphysical possibilities is to use examples and compare them with other kinds of possibilities, like physical possibilities and logical possibilities. The different kinds of possibility will be defined more precisely in the next chapter, but here we can get an initial idea of these kinds and see where metaphysical possibilities stand in this picture.

It is generally agreed that, if we define physical possibility as what is compatible with the laws of nature, metaphysical possibility is wider than physical possibility, in the sense that what is metaphysically possible does not need to be compatible with these laws, and not all that is metaphysically possible is also physically possible. At the same time though, if we define logical possibility as what is compatible with the laws of logic, it is thought that metaphysical possibility is narrower than logical possibility, in the sense that not all that is logically possible is also metaphysically possible. While metaphysical possibility includes and exceeds physical possibility, logical possibility includes and exceeds metaphysical possibility.

Consider the following counterfactuals:

(10) If a ginger cat had kittens, they would be ginger

(11) If a ginger cat laid eggs, ginger kittens would come out from them

(12) If George were drinking water but not H_2O, he would get poisoned

While (10) involves a metaphysical possibility that is also a physical possibility, because it is physically possible for a ginger cat to have kittens, (11) involves a metaphysical possibility that is not a physical possibility as well, because it is not physically possible for a ginger (or any) cat to lay eggs. On the contrary, (12) involves a logical possibility that is not also a metaphysical possibility, because it is not metaphysically possible for George (or anybody) to drink water that is not H_2O.

We will put on a side the cases like (12), and so those examples of logical possibility exceeding metaphysical possibility. Even if logical possibility is a more comprehensive notion than metaphysical possibility, it is in many ways too "liberal" to be practically helpful in the definition of counterfactuals, because it is only restricted by the rules of logic. The

counterfactuals on which we want to focus, instead, are somehow more "tangible", more linked to how things stand (or do not stand) in the world. Therefore, in this work, we prefer to relate counterfactuals specifically to metaphysical modality, leaving that part of logical modality that exceeds metaphysical modality out of the discussion.

Putting on a side case (12) then, we can see that while (10) can be easily evaluated by considering the laws of nature (e.g., the laws of genetics concerning cats), it is more difficult to determine how to evaluate (11) because these laws do not seem to apply. However, we would not consider that we cannot evaluate it, or that we are compelled to evaluate it as false only because it is against the laws of the actual world that cats lay eggs. Indeed, there could be an argument that (11) might well be true.

As we said in section 3, many counterfactuals are contrary to facts, in the sense that their antecedents involve situations that are in contrast with, or are different from, what is the case in the actual world, and their contents relate to what would result from those differences. These differences usually imply a violation of the laws of the actual world. Therefore, counterfactuals seem particularly prone to involve metaphysical possibilities, because these possibilities do not need to be compatible with the laws of the actual world, and so can include those situations that are in contrast with the actual world. However, the fact that the antecedent of these counterfactuals is against the facts of the actual world is not taken to be a reason to consider all of them false, because it is only after the evaluation of the consequent given such an antecedent that we can tell whether they are true or false. What is needed for evaluating these counterfactuals then is an account of metaphysical possibility, to explain how their antecedent can be against the facts.

As a matter of fact, to account for metaphysical possibilities is one of the main challenges for modal discourse, because they cannot be evaluated in terms of the actual world in the same direct and natural way as physical possibilities. This is why an account of metaphysical modality usually requires a complex metaphysical background, that can explain how we can deal with possibilities exceeding the laws of the actual world. The development of such a background is the job of modal metaphysics, the branch of metaphysics that aims to explain modality. One of the prominent accounts of modal metaphysics that was also applied to counterfactuals is the possible world account of modality, where metaphysical modality is explained in terms of quantification over metaphysically possible worlds. However, here I want

to pursue another account of modal metaphysics, the potentiality account of modality, which also applies to counterfactuals.

A more detailed discussion of modality will happen in the next chapter, so for now we can simply conclude that counterfactuals do include cases in which the antecedents are metaphysically possible, but not actually true nor physically possible, and therefore it seems appropriate to include them in modal discourse and connect them to an account of modal metaphysics.

4.2. Possibilities in Time

In addition to dealing with metaphysical possibilities, counterfactuals also involve possibilities in time: sometimes they talk about present or future possibilities, other times they talk about possibilities in the past. Take the next two examples:

(13) If the glass were struck, it would break

(14) If the glass had been struck, it would have broken

Taking (13) first, intuitively the counterfactual is telling us what *could happen* to the glass either in the present or in the future. The glass has not been struck yet, and we do not know as a fact that it is being struck at the present moment or that it will be struck at some future moment. Within this counterfactual we conceive the possibility for the glass to be struck at any moment from now and imagine what would happen. We *do not know* that it will be struck, but we can ponder on the possibility for this to happen. The possibility involved here is open towards the future, and the counterfactual can work like some sort of prediction. This is the kind of counterfactual you would use in designing an experiment or in deciding what to do in the future, thinking of what the outcomes of our decisions could be.

Talking about (14) instead, the counterfactual expresses something that *could have happened* to the glass in the past. It seems odd that we can talk about possibility when dealing with something closed and determined as the past, but this is what past counterfactuals do, thinking of other ways things could have been. This is particularly helpful when we judge our past decisions, wondering whether having acted differently we would have produced another result: this is how we learn from past mistakes and successes.

Nevertheless, in talking about past possibilities, counterfactuals can take two approaches. In some cases, past counterfactuals consider the possibility of the happening of an event that we know did not happen and its hypothetical results, like in our previous example (6):

(6) If Napoleon had not been defeated in Waterloo, he would not have been exiled in Saint Helena

In the case of these counterfactuals, which are usually called 'historical counterfactuals', our knowledge of the past makes us contemplate an alternative to this past in which the things that we know happened, or did not happen, have the opposite outcome.

In other cases, counterfactuals simply refer to possibilities in the past. Let us consider (14) again:

(14) If the glass had been struck, it would have broken

In this case, we are considering the possibility that the glass was struck at some moment in the past with respect to the counterfactual. Even if, intuitively, we might do so, in (14) we do not need to suppose that the glass was not struck in the past, as we might not know if it was struck or not, but with the counterfactual we consider the possibility for this to have happened.

We will come back on the difference between past counterfactuals in general and historical counterfactuals, but for the moment we can conclude this section by saying that, in producing truth conditions for counterfactuals, we need to include a reference to both metaphysical possibilities and possibilities in time.

It is quite clear then that, in looking for an account of counterfactuals, we cannot overlook a discussion of how to account for possibility in general. This is why counterfactuals have been included in modal discourse, which usually includes necessity and possibility. Inserting counterfactuals in the category of modality is almost inevitable and has many advantages, but it also ties such an account of counterfactuals to the problems coming from an account of possibility and necessity. The connection to modality is what creates most of the complications surrounding counterfactuals, because modality itself creates many difficulties both in metaphysics and epistemology. The problem of how to deal with possibility and how to relate counterfactuals to modality will be matter of the next chapter.

For the moment, we will stay on counterfactuals, offering a small list of the most important types of them.

5. Kinds of Counterfactuals

When, at the beginning of this chapter, we talked about the counterfactual structure, we already pointed out that we do not consider to be counterfactuals only those presenting the "If *it were the case that* A, then *it would be the case that* C" form. In what follows, we will introduce the main kinds of counterfactuals that need to be taken under consideration. This list is based on the types of verbs we can find in different counterfactuals, showing how a different verbal tense or modal verb can influence our approach to the counterfactual.

The discussion of the relation between might/could-counterfactuals and would-counterfactuals is of crucial importance for Vetter's potentiality account of counterfactuals so we will discuss it more in depth in the next section.

5.1. Would-counterfactuals

These are the basic kind of counterfactuals, and we can take as example (13) above:

(13) If the glass were struck, it would break

In general, these counterfactuals have verbs in the past tense like 'had', 'did' and 'were' ("If I had a car…", "If John did not play football…", "If I were you…", etc.), or the construction 'were to' ("If I were to go shopping…", etc.) in the antecedent, and verbs preceded by 'would' in the consequent ("…then I would drive to work", "…then John would not have so many trophies", "…I would not be so upset", "…then I would buy some milk", etc.). In the course of this work, we will refer to these counterfactuals when talking about 'would-counterfactuals'.

These are the sort of counterfactuals we described in section 4.2 as dealing with present and future possibilities. In particular, those would-counterfactuals using the construction 'were to', such as "If Jane were to cook dinner, we would eat some amazing food", seem to suggest very strongly an interpretation towards the future. However, this temporal interpretation must not mislead us. These counterfactuals do not need to be so clearly set in a present or future time: we can have examples that are almost "timeless" like "If there were unicorns in the world,

we would feed them with carrots", in which we are dealing with a possibility not clearly defined in terms of time.

The use of 'were to' in counterfactuals requires a short explanation: it adds either a stronger future reading to the counterfactual, or an indication that the antecedent is unlikely, depending on the situation: "If I were to go shopping…" can suggest both that the action will happen at some future time and that it is not likely that I will go shopping, or that I have no intention to go shopping. However, the use of 'were to' does not seem to imply a variation in the truth conditions of a counterfactual, so in this work it will be treated as an equal alternative to the standard form, whose difference pertains only to the pragmatics of language.

Would-counterfactuals are usually the starting point of an account of counterfactuals, which is normally built by finding truth conditions for this kind of counterfactuals first. Barbara Vetter's account represents an exception, as she starts from could-counterfactuals, as we will see in later chapters.

5.2. Past-counterfactuals

These are the counterfactuals dealing with past possibilities, so we can take as example (14) above:

(14) If the glass had been struck, it would have broken

In these counterfactuals, the verbs in the antecedent are in the past perfect tense like 'had been' and 'had done' ("If I had drunk one more glass…", "If John had not been so absent-minded…", etc.) and the verbs in the consequent are preceded by 'would have' (…then I would have been sick", "…then John would not have missed his bus", etc.). In general, we will call these counterfactuals 'past-counterfactuals' across the thesis.

As discussed in section 4.2 above, these are the counterfactuals involving past possibilities. As said in that section, this form is the one mainly associated to the 'against-the-facts' interpretation of counterfactuals as it tends to report the possibility for something we know it happened in the past to have been different. When this is the case, a past-counterfactual is said to be a 'historical counterfactual': these are the counterfactuals that are used to talk about how the past could have been different and usually involve the changing of some well-known past facts. However, we also said that this is not the necessary interpretation of this form, as it can also just consider a possibility in the past without knowing the facts about it.

Again, though, the temporal interpretation might not be so strict in some cases: the example "If there had been unicorns in the world, we would have fed them with carrots" seems equally "timeless" as the would-counterfactual version above.

Any account of counterfactuals aiming to be exhaustive must include an account of past-counterfactuals, both in general and as historical counterfactuals. Therefore, the framework in which this account is inserted must offer an account of past possibilities.

5.3. Might/could-counterfactuals

These are counterfactuals presenting a similar structure to would-counterfactuals, but including 'might' and 'could' in the place of 'would', like:

(15) If the glass were struck, it might break

(16) If the glass were struck, it could break

These counterfactuals still have the verbs in the antecedent in the past tense, like 'had', 'did' and 'were', or in the 'were to' form, ("If I had a car…, "If I were to go shopping…, etc.) but instead of having the verb in the consequent preceded by 'would', the verb is preceded by the modal verbs 'might' or 'could', (…then I could drive to work", "…then I might buy some sweets", etc.). These will be respectively called 'could-counterfactuals' and 'might-counterfactuals' from now on. The difference between might- and could-counterfactuals is usually regarded as less crucial than the difference between these two and would-counterfactuals, but it might be worth just make some clarifications.

Considering the two modal predicates 'may' and 'can', from which the two counterfactual variants 'might' and 'could' are derived, we cannot deny that there is a difference between how they are used. On one side, 'may' is used in general to express something on which the speaker is uncertain. On the other side, 'can' is mostly used to express an ability, or a concession, or the possibility for something to happen.

To explain this difference, we can introduce a distinction used by Vetter, and that will be discussed in depth in later chapters, between 'dynamic modality' and 'epistemic modality': dynamic modality is focusing on what could happen depending on how things really are – therefore it depends on "the reality", while epistemic modality is focusing on what is compatible with our knowledge – therefore it depends on our *knowledge* of "the reality". 'Can', as an

expression of the possibility for something to happen, is generally taken to express dynamic possibility, while 'may', as an expression of uncertainty, is generally taken to express epistemic possibility.

However, despite these strong intuitions about the two, 'may' and 'can' are often considered interchangeable. This is because there is an intrinsic and natural *ambiguity* in modal expressions: "John can be home in an hour" can be read not only as the ability for John to be home in an hour, but also as the epistemic possibility expressed by a speaker that John will be home in an hour, and so as we would normally read "John may be home in an hour". It is a generally accepted feature of modal expressions that the same modal words can express different kinds of modality, depending on the reading we attribute to them.

Therefore, 'can', despite its default interpretation as a dynamic modal, can be read as expressing epistemic possibility, and 'may', despite its default interpretation as an epistemic modal, can be read as expressing dynamic possibility. Even if we agree that dynamic modality and epistemic modality must be treated differently, 'may' does not need to be read as an epistemic modal in all its instances, as 'can' does not need to be read as a dynamic modal on every occasion, so that the two terms can actually be used interchangeably if they have the same reading.

The same is valid for 'could' and 'might', so that might-counterfactuals can be read as could-counterfactuals and vice-versa. Therefore, the difference between might- and could-counterfactuals is not relevant *per se*, but only depending on the reading we attribute to 'might' and 'could'. This clarification is quite important for what follows because it will allow Barbara Vetter to group together could- and might-counterfactuals, which will be a crucial step for developing her account of would-counterfactuals, as we will see.

Before discussing the all-important difference between might/could-counterfactuals and would-counterfactuals, we should just point out that there are both might- and could- past-counterfactuals:

(17) If the glass had been struck, it could have broken

(18) If the glass had been struck, it might have broken

These counterfactuals are the past versions of might/could-counterfactuals so we can put together what just said about these with what said about past-counterfactuals. For what concerns past possibilities, we can refer on what said in 5.2, as also these two can either simply

refer to some possibilities in the past or be historical counterfactuals. Also, the fact that we are talking about past possibilities does not seem to influence the fact that 'could' and 'might' can both be read as expressions of epistemic or dynamic modality so that the difference in intuitions is still contrasted by the possibility of ambiguous readings. In what follows, these past-counterfactuals will not be discussed directly, as most of the explanation about them can be gathered by putting together what we will say about might/could-counterfactuals and past-counterfactuals.

6. Would- vs Might/could-counterfactuals

After discussing the main types of counterfactuals, it is important to focus on the difference between would-counterfactuals and might/could-counterfactuals.[7] To explain this difference, I will first address it following our intuitions on some examples. I will use might-counterfactuals rather than could-counterfactuals as they are the ones that are traditionally contrasted with would-counterfactuals. Consider the following pair:

(19) If they had not hired Jones, they might have hired you

(20) If they had not hired Jones, they would have hired you[8]

Trying to give this an intuitive reading, we could say that a speaker tends to use (19) when there is a possibility that they would have hired you if they had not hired Jones, but the speaker seems to intend that this is still just a mere possibility, and they could still not have hired you in the same way. In (20) on the contrary, the speaker seems to intend that you would have been surely hired if Jones had not been, as if it is more than just a mere possibility.

The difference between 'would' and 'might' has been treated differently in the literature: David Lewis (1973a, p. 1-2) links two different logical symbol to would-counterfactuals and

[7] Some parts of this section have previously appeared in my article (Casini, 2022).

[8] Note that here I am using past-counterfactuals as examples, while in most cases across the thesis Lewis' translations will be applied to present-/future-directed counterfactuals. While there is no formal difference in applying the translations to past or present/future counterfactuals, the past examples presented here have the heuristic intent to offer a particularly clear picture of the difference and the relation between might- and would-counterfactuals.

might-counterfactuals, while Robert Stalnaker (1981, pp. 99-101; 1987, pp. 142-6.) supports a view that there is no logical difference between the two.

According to Stalnaker, might- and would-counterfactuals do not require a different logical definition. He suggests that the difference between 'would' and 'might' lays only on the fact that 'might' expresses epistemic possibility, so that might-counterfactuals simply are a representation of the speaker's epistemic status towards the consequent given the antecedent. Therefore, in his view, (19) should only mean "It is epistemically possible that if they had not hired Jones, then they would have hired you".

According to Lewis, instead, as we define possibility using necessity and vice versa, we can use would-counterfactuals to define might-counterfactuals and vice versa. In *Counterfactuals* (1973a, p. 21), he presents the following "translations" between might- and would-counterfactuals:

I. "If it were that *P*, it would be that *Q*" can be translated as "It is not the case that if it were that *P*, it might be that not *Q*".

II. "If it were that *P*, it might be that *Q*" can be translated as "It is not the case that if it were that *P*, it would be that not *Q*".

Therefore, in Lewis' view, (19) and (20) can be translated as follow:

(19*) If they had not hired Jones, they might have hired you = It is not the case that if they had not hired Jones, then they still would not have hired you

(20*) If they had not hired Jones, they would have hired you = It is not the case that if they had not hired Jones, then they might not have hired you

From these translations we can gather that Lewis' position seems to suggest too that the consequent of would-counterfactuals, on the assumption of the antecedent, looks in a way 'more fixed' or 'more determined' than the consequent of might-counterfactuals. The idea is that when we use a counterfactual, we first consider the possibility for the antecedent to be the case, but then the use of 'would' seems to indicate more strongly that, given the antecedent, the consequent should be the outcome, while the use of 'might/could' seems to indicate that, given the antecedent, the consequent is only one of the possible outcomes. Here with "possibility for the antecedent to be the case" we do not mean to exclude counterfactuals with impossible antecedents, rather we mean that we are taking the antecedent to be the case for the sake of

the evaluation of the counterfactual, whether this antecedent is possible or not, as we explained earlier.

If we consider the counterfactual structure to be already an indicator that we are considering the possibility for the antecedent to happen (even when it is impossible), the use of 'might' adds an extra indicator, that the consequent is only *possible given the antecedent*. More precisely, the counterfactual structure indicates that we are considering the possibility for the antecedent to be the case, but with 'would' this structure indicates that the consequent *must be* the result of these circumstances, while with 'might' it indicates that the consequent is only *possible* under these circumstances.

To show this even further, we can consider the following formal transcriptions of Lewis' translations:

(WT) $(p \Box\!\!\rightarrow q) =_{df} \sim(p \Diamond\!\!\rightarrow \sim q)$

(MT) $(p \Diamond\!\!\rightarrow q) =_{df} \sim(p \Box\!\!\rightarrow \sim q)$

Where '$\Box\!\!\rightarrow$' is the symbol for would-counterfactuals, and '$\Diamond\!\!\rightarrow$' is the symbol for might-counterfactuals (with '~' being the symbol for negation). It is immediately noticeable how these translations seem reflect the principle of interdefinability of the possibility '◊' and necessity '□' operators:

(N/P) $\Box p =_{df} \sim\Diamond (\sim p)$

(P/N) $\Diamond p =_{df} \sim\Box(\sim p)$

The similarities between the two (WT) and (MT) and the two (N/P) and (P/N) do not stop here for Lewis, since the definitions of '$\Box\!\!\rightarrow$' and '$\Diamond\!\!\rightarrow$' that he provides follow a similar pattern of the definitions of '◊' and '□' in terms of possible worlds as well. As we will explain in the next chapter, necessity ('□') is defined as truth at all possible worlds and possibility ('◊') is defined as truth in at least one possible world, and Lewis' definitions of the truth conditions for would- and might-counterfactuals can be presented in possible world terms as follow:

(a) A counterfactual $p \Box\!\!\rightarrow q$ is true at world *w* iff all the possible worlds most similar to *w* at which *p* is true are worlds at which *q* is true.

(b) A counterfactual $p \Diamond\!\!\rightarrow q$ is true at world w iff there are some possible worlds sufficiently similar to w at which p is true that are worlds at which q is true.[9]

Despite these similarities, it is worth noting that we cannot explain the difference between would- and might/could-counterfactuals explicitly in terms of necessity and possibility, saying that the consequent of a would-counterfactual is necessary, while the consequent of a might/could-counterfactual is possible. Even if it seems a straightforward and natural interpretation of Lewis' point, it is always given the possibility for the antecedent to be the case that then we evaluate whether the consequent is necessary or possible. Indeed, the "necessity" of the consequent here is not 'necessity' as normally intended in possible world terms either: if necessity is truth at all worlds, the "necessity" of the consequent of a would-counterfactual is truth only at all those worlds that are most similar to the actual world where the antecedent is true. Same for the "possibility" of the consequent. Therefore, saying that the consequent is necessary or possible should always be under the clause '*given the antecedent*'.

Even if in explaining the difference between might/could- and would-counterfactuals we should be careful, we can summarize it by saying that 'might' and 'could' (in their dynamic reading) both indicate that the consequent is *possible given the antecedent*, in opposition to 'would' which indicates that the consequent is *necessary given the antecedent*. The relevance of this distinction will become clear later.

Coming back to 'could' and 'might', we will support the idea that they can be grouped together, so that what we said on the difference between 'would' and 'might' is also valid for the difference between 'would' and 'could'. Therefore, we can take 'could' as an alternative to 'might', in the sense that they both indicate possibility relative to the consequent given the antecedent in opposition to 'would', that indicates necessity.

Naturally, the distinction between 'would' and 'might' is quite controversial. The duality thesis between would- and might-counterfactuals supported by Lewis has not been so straight-forwardly accepted by everyone. Other than Stalnaker (1981, 1987), others have been sceptical (e.g., De Rose 1994, 1999). The debate on the relationship between would- and might-

[9] The interdefinability principle between necessity and possibility as well as Lewis' possible world account of counterfactuals will be discussed and clarified later in this work, but it seemed useful introducing some of these technicalities here to show more clearly Lewis' position on might-counterfactuals.

counterfactuals arises from the natural ambiguity of modal expressions, and there is still a lot of controversy on how might-counterfactuals should be interpreted (see Eagle 2007; Ferguson 1987). I will not enter in the heat of this debate here as I do not wish to investigate this issue in too much depth.

However, in this thesis I will support Lewis' idea of interdefinability between would- and might/could-counterfactuals, as this is a necessary requirement for formulating an account of would-counterfactuals in terms of potentiality. We will come back on the relation between these counterfactuals, and we will see the importance of the interdefinability principle for Vetter.

7. Conclusion of Chapter I

This concludes our initial investigation on counterfactuals. We discussed the relationship between counterfactuals and conditionals in general in section 1. In section 2, we compared counterfactuals to material and indicative conditionals. We explained the difference between the terms 'counterfactual' and 'subjunctive conditional' in section 3. Section 4 was dedicated to the relation between counterfactuals and possibility. Finally, section 5 listed the kinds of counterfactuals that an effective account needs to take under consideration and section 6 investigated the difference between would- and might/could-counterfactuals.

Our aim in this first chapter was to stress few key points. First, although counterfactuals are conditionals, they are not like the others, both because of their structure and because the accounts of other conditionals are not satisfactory in explaining them. Second, counterfactuals are intrinsically sensitive to possibilities, and in some cases to metaphysical possibilities that are not easily explained by appealing to the actual world, which therefore need to be accounted for in talking about counterfactuals. Third, counterfactuals also involve past possibilities, so these possibilities as well must be considered in an account of counterfactuals. Fourth, counterfactuals have different shapes, in particular as would- and might/could-counterfactuals, and we need to account for all of them.

Now it is time to move our inquiry to the crucial relation between counterfactuals and modality.

CHAPTER II

COUNTERFACTUALS AND MODALITY

The aim of this chapter is to understand better the relation between counterfactuals and modality, and to explain some of the reasons why counterfactuals have been included in modal discourse.

To do this, now that we know what counterfactuals are, we need to investigate what modality is. This not only can shed light on how we conceive possibility and necessity, but also show why it is appealing to include counterfactuals in modal discourse. The main reason for the inclusion of counterfactuals within modality is the fact that they are sensitive to metaphysical possibilities that are not so easily explained by appealing to the actual world. Modality concerns all kinds of possibility, including metaphysical possibilities, and if counterfactuals involve these possibilities, it seems natural to include them in modal discourse. As a matter of fact, the various distinction between different kinds of possibility and necessity within modality can prove very helpful in grasping what metaphysical possibilities are.

The discussion of metaphysical possibilities, in particular those that exceed physical possibilities, will offer some background to the possible world account of counterfactuals and will clarify some of the motivations that led towards developing such account. One motivation is that, since for finding the truth conditions of counterfactuals we need to appeal to metaphysical possibilities, one way to account for this sort of possibilities seems not to be based on the actual world alone, but also on certain non-actual worlds. However, this account is not the one preferred in this thesis, which wants to pursue an account of counterfactuals based on potentiality.

This chapter will suggest that, for formulating an account of counterfactuals, it is necessary to provide an account of metaphysical possibility, which requires the introduction of a metaphysical background that can explain how certain possibilities can exceed the actual world. The definition of such a background is the main aim of the so-called 'modal metaphysics': the branch of metaphysics focusing on modal discourse. In this thesis, we will focus on one account of modal metaphysics: Barbara Vetter's potentiality account of modality. However, we will not need to discuss this account in its entirety in too many details, because what matters to us is how it applies to counterfactuals.

The structure of the chapter will be the following. The first section will introduce the concept of 'modality', while the second will discuss the different ways to conceive modality and possibility, introducing and explaining the notion of 'metaphysical possibility' through different classifications of modality. The third section will describe what problems arise within modal discourse, focusing in particular on the Metaphysical Problem of Possibility. The fourth section will discuss where counterfactuals stand in the modal framework and how the connection to modality can influence an account of counterfactuals. The conclusion of the chapter will lay the ground for the introduction of Vetter's potentiality account of modality.

1. Introduction to Modality

In the previous chapter, we introduced the idea that counterfactuals are considered part of modality. A reason in favour of this idea is that counterfactuals involve metaphysical possibility. In this chapter, we will explore more in depth what modality is and the different ways in which possibility can be conceived, trying to offer a better account of the relationship between counterfactuals, possibility, and modality. This will be preparatory for a better understanding of why counterfactuals seem to need an account of modal metaphysics.

Nowadays, when we use the term 'modality', we immediately associate it with necessity and possibility, to such an extent that we cannot really think of these terms as separated. This is the result of a long story of how the concept of modality evolved.

Originally, the term 'modality' was used to indicate the *mode of existence* of certain entities (Borghini, 2016, p. 10). The idea was that, when a modal expression is applied to something, this tells us about the way in which this something exists, so that expressions like 'necessarily' or 'possibly' tell us that the thing to which they apply exists in a necessary way or

in a possible way. Within this framework, modalities were regrouped in different categories, corresponding to different *modes*, as effectively listed by von Wright (1951, p. 2):

alethic	**epistemic**	**deontic**	**existential**
necessary	verified	obligatory	universal
possible		permitted	existing
contingent	undecided	indifferent	
impossible	falsified	forbidden	empty

In von Wright'a definitions, alethic modalities are the modes of truth, epistemic modalities are the modes of knowing, deontic modalities are the modes of obligation and existential modalities are the modes of existence[10] (von Wright, 1951, pp. 1-2).

Both the idea of *modes of existence* and von Wright categorization are not considered instrumental for a definition of modality anymore. First, it is now recognised that modality does not concern only ascriptions of the ways in which an object exists, but also ascriptions of how certain properties belong to objects: we do not just say "*x* necessarily exists" but also "*x* necessarily has the property of being *y*". Second, between the kinds of modalities listed by von Wright only the alethic and deontic are considered relevant to the discussion of modality in a modern sense. As a matter of fact, when we talk about 'modality' these days, we mainly acknowledge necessity and possibility to be part of it[11], both in their alethic sense of necessarily true or possibly true, and in their deontic sense of necessary/obligatory by laws or moral constraint and possible/permitted by laws or moral constraint. 'Modality' then is nowadays mostly used to indicate only these two sets of modes, rather than the various sets presented by von Wright.

[10] Note that this is different from the 'mode of existence' in Borghini's terms, in which this is very generally 'the way some entity exists'. When von Wright talks about existential modalities, he refers to the universal and existential quantifiers of quantified logic.

[11] This is not entirely true, as the category of modality is now believed to include many other 'modalities', like counterfactuals, essences, and dispositions, as we will explain later. However, for the moment we will focus on necessity and possibility.

Note that with this we do not mean that the other 'modes' are not object of discussion and study in philosophy. For example, what von Wright called 'epistemic modality' is widely discussed in epistemology, even if it does not seem to be crucial for the semantics and logic of "modality" strictly speaking in a modern sense. This is because 'epistemic modality' in von Wright's terms does not correspond to the category of 'epistemic modality' that we will discuss later: the former concerns verification and falsification, while the latter concerns compatibility with one's knowledge.

Between the two, we will put to one a side deontic modality, letting our focus being on alethic modality, which includes necessity and possibility as we normally understand them when we use sentences like "*x* is necessary" or "*x* is possible". Coming back to von Wright's interpretation, alethic modalities are thought to be the modalities concerning the ways in which something is true (Borghini, 2016, p. 10). Nowadays, this seems not to reflect entirely our intuitions on modalities, as when we say "*x* is necessarily/possibly true" we seem to intend more than one thing, as we are going to see. Even if the original idea that alethic modalities are modes of truth is somehow outdated and needs a deeper analysis, it was correct in stating that truth is essential to the definition of alethic modality. Indeed, both alethic necessity and alethic possibility can be defined in terms of truth. It is quite intuitive that if something is true, it is possible. Equally intuitive is that if something is necessary, it is true. Therefore, we can define alethic necessity as the sort of necessity that implies truth, and alethic possibility as the sort of possibility that is implied by truth.

This brings us to one of the interesting features of alethic modalities that we have already introduced in the last chapter: they are *interdefinable*. If we consider that we apply these modalities to propositions, then we could formulate the following interdefinitions:

- A proposition is necessary when its negation is not possible.
- A proposition is possible when its negation is not necessary.
- A proposition is contingent when both the proposition and its negation are possible.
- A proposition is impossible when its negation is necessary.

The interdefinability of alethic modalities is central for the discussion of modality. The fact that we can define necessity in terms of possibility and vice versa on the one hand implies that, if we take one of necessity and possibility to be primitive, we only need one of the two

concept to be primitive, as the other can be defined in terms of the first; on the other hand it implies that, if we do not consider necessity and possibility to be primitive, once we got a definition for one of the two concepts, we can work out the definition of the other.

Nowadays, alethic modality is said to be of two kinds: we can distinguish between modality *de dicto* and modality *de re*. Modality *de re* concerns how an object possesses certain properties: "John is necessarily human" is a *de re* proposition, saying "something about the way in which a particular object instantiates a property" (Melia, 2003, pp. 26-27). Modality *de dicto* instead, concerns propositions rather than objects: "Necessarily, all bachelors are unmarried" is a *de dicto* proposition, which "ascribe[s] a necessity to the proposition as a whole rather than any particular individual or object" (Melia, 2003, p. 27). Note that this metaphysical reading of the distinction usually reflects, but must be clearly distinguished from, a purely logical reading of it, where the contrast is between *de re* modality as the case where the quantifier takes scope over the modal operator ("Some people are necessarily bald") and *de dicto* modality as the case where the modal operator takes scope over the quantifier ("Necessarily, some people are bald").

The distinction between *de dicto* and *de re* modality is not the only one introduced to offer a better account of modality. If we could consider this a sort of "macro" distinction between two ways in which modalities apply to entities (propositions vs objects), there are also other distinctions, depending on how we define necessity and possibility once they are applied to entities, being these objects or propositions.

It is quite intuitive that necessity and possibility do not correspond to a unique definition. Borrowing a good example from Melia (2003, p. 16), in saying "It is possible for me to speak French" we can mean two different things: either we mean that I have studied French, I learnt it and I understand it, so that I can speak the language, or we mean that, because I am a human being, and human beings can learn different languages, it is in my power to speak French, if I want to learn it. In the first case, we mean that I have the *ability* to speak French; in the second, we mean that, even if I do not have this ability at the moment, it is possible for me to acquire it. The term 'possible' here is not univocal, suggesting that there is more than one

sense in which we can use it.[12] In the next section, we are going to investigate some of these senses.

2. Classifying Modality

Modality has been classified in many ways across the literature. In what follows, we will consider two of them.

The first one, which we will present through a scheme that can be found in John Divers' book *Possible Worlds* (2002), is more traditional and defines the different ways in which a proposition can be said to be 'possibly true' or 'necessarily true' based on a possible world picture of modality. In short, the kinds of modality that are introduced are taken to correspond to different kinds of possible worlds that are selected when evaluating a proposition. Even if the intent of this thesis is to investigate how well a metaphysical background of modality based on potentiality, and not on possible worlds, can relate to an account of counterfactuals, it still sometimes can be more intuitive and easier to use possible worlds to explain many of the concepts involved in modal discourse. However, I want to stress that our appeal to possible worlds as a heuristic tool to facilitate my explanations does not imply my support to a possible world metaphysics. Even if I do recognise their importance and their utility in modal discourse, I do not consider possible worlds a satisfactory metaphysical background.

The second classification is the one adopted by Barbara Vetter in her book (2015), and we will discuss it because it is instrumental to understand some aspects of her account of counterfactuals. This second classification stresses different distinctions between modalities than the traditional one, but I will not take a position on which classification is the best, as I do not need to. The two are not incompatible and are both helpful: the traditional classification helps understanding the concept of metaphysical modality, while the other is important in understanding certain positions of Vetter concerning her approach to potentialities, as we will see.

[12] Note that this is the view defended by Vetter in her book (2015), but the orthodox view in semantics, like the one supported by Angelika Kratzer (1989, 2012) is that we have here a case of context-sensitivity, not ambiguity.

It is worth noting that classifying modality is not entirely straightforward and that different philosophers seem to adopt different classifications, attributing to the same terms different meaning. The two versions presented here are nothing but a portion of the many possible ways to categorize modality. They can both be controversial in some ways, but they are instrumental to our discussion, and so we will not engage the possible criticism in this work.

2.1. Modality and Possible Worlds

In this section, we will see how possible worlds can be used to define modality. Even if I will ultimately abandon a possible world account of modality in favour of a potentiality-based one, I still think that what follows can help to understand what modality, and its categories, are. The following presentation is based on John Divers' book *Possible Worlds* (2002).

Divers defines modality, or modal discourse, as the study of concepts like necessity and possibility and he suggests that we can identify four "cases" of modality, which are the basic modal concepts that modality aims to explain and that we have seen above: possibility, necessity, impossibility, and contingency (2002, p. 3). Along with these, Divers lists seven "kinds" of modality, which are varieties of "constrained" modalities that can all be included under the label 'modality': logical modality, analytical modality, metaphysical modality, physical modality, epistemic modality, doxastic modality, and deontic modality (2002, p. 4). The utility of possible worlds in modal discourse lies in the fact that they are considered a highly effective tool for explaining both the cases and the kinds of modality.

On the one hand, possible worlds are invoked to define the cases of modalities by offering definitions of each that are based on quantifications over possible worlds, following the scheme below:

- Possibility is truth in at least one possible world
- Necessity is truth in all possible worlds
- Impossibility is truth in no possible world
- Contingency is truth in some but not all possible worlds

The cases of modality are traditionally considered interdefinable, as we know. Possibility is what is not impossible and is implied by necessity and contingency. Necessity is what is not impossible nor contingent and requires possibility. Impossibility is what is not possible, nor necessary, nor contingent. Contingency is what is not impossible nor necessary

and requires possibility (see Divers, 2002, p. 3-4). Possible worlds are very useful in explaining this interdefinability as well because, following the principles of quantification, we can see that: necessity implies possibility because what is true in all possible worlds is true in some; impossibility excludes necessity because what is true in no possible worlds is not true in all possible worlds; contingency requires possibility because what is true in some but not all possible world is obviously true in some; etc (see Divers, 2002, p. 5).

On the other hand, possible worlds are helpful in accounting for the kinds of modalities, which are most easily identified if we consider the different kinds of impossibility, corresponding to "different kinds of consideration that exclude something from the realm of possibility" (Divers, 2002, p. 4).

These excluding considerations can have different nature, so that they exclude different impossibilities and therefore define the "limits" of the different kinds of modality.

Logical modality. When these excluding considerations are logical (e.g., no proposition can be both true and false), they exclude logical impossibility and so define the limit of logical modality.

Analytical modality. When the considerations depend on the meaning of words (e.g., no bachelor can be married), they exclude analytical impossibility and so define analytical modality.

Metaphysical modality. when the considerations depend on "the natures and identity conditions of things" (Ibid.) (e.g., no one can drink water without drinking H_2O) they exclude metaphysical impossibility and so define metaphysical modality.

Physical modality. When the considerations depend on the laws of nature (e.g., nothing with a mass can travel faster than light), they exclude physical impossibility and so define physical modality.

Epistemic modality. When the considerations depend on what is known, excluding what goes against what one knows (e.g., no cat can lay eggs because I know they are mammals and mammals do not lay eggs), they exclude epistemic impossibility and so define the limit epistemic modality.

Doxastic modality. When the considerations depend on what is believed, excluding what goes against what one believes (e.g., my daughter cannot be at the playground

because I believe she is at school, where I left her), they exclude doxastic impossibility and so define the limit of doxastic modality

Deontic modality. When the considerations depend on certain rules and norms (e.g., a driver cannot pass with a red traffic light) they exclude deontic impossibility and so define the limit of deontic modality.

These kinds of modality stand traditionally in various interrelations, with logical modality entailing metaphysical modality, but not vice versa, and metaphysical modality entailing physical modality, but not vice versa, following this scheme: all metaphysical possibilities are logical possibilities, but not all logical possibilities are metaphysical; all physical possibilities are metaphysical possibilities, but not all metaphysical possibilities are physical. Once again, possible worlds can account for these interrelations as well, as we are going to see.

If possible worlds can explain the different cases of modality in terms of quantification over possible worlds, they can explain the different kinds of modality by applying *constraints* upon this quantification, in the following way:

> "We begin with the idea of the totality of the possible worlds across which all of the genuine possibilities (and no impossibilities) are represented. One of these possible worlds – the actual world – is special, closer to our hearts and distinguished somehow for the others that are 'merely' possible. Given the conception of a plurality of possible worlds, a modality of a given case and kind is characterized in terms of what is the case at a specific range of possible worlds" (Divers, 2002, p. 5)

In this way we can identify some M-possible worlds (the specific ranges of possible worlds) as those among the plurality of all genuine possible worlds that conform to the set of M-constrains (the excluding considerations identified above) (*Ibid*.). Therefore, we will have logically possible worlds, that are those complying to the laws of logic; analytically possible worlds, that are those complying to meanings; metaphysically possible worlds, that are those complying to the identity and the nature of things; physically possible worlds that are those complying to the laws of nature; and so on.

Following this then, we can say that something is M-possible if it is true at some M-possible world (e.g., logically possible if it is true at some logically possible world); something is M-impossible if it is not true at any M-possible world; something is M-necessary if it is true

at all M-possible worlds and something is M-contingent if it is true at some but not all M-possible worlds (*Ibid*.).

Such characterisation, based on the definition of constraints upon the quantification over possible worlds, makes clearer the interrelation between the kinds of modality. The application of constraints on the selection of worlds over which quantify shows, for example, that metaphysical possibility entails logical possibility but not vice versa – with all metaphysically possible worlds being logically possible worlds, but not all logically possible worlds being metaphysically possible worlds – because the constraints applied to select only the metaphysically possible worlds will produce a smaller collection of worlds than the constraints selecting the logically possible worlds, so that the metaphysically possible worlds are a subset of the logically possible worlds.

It is controversial whether any of the collections of M-possible worlds should correspond to the collection of all the genuine possible worlds, so that, for example, we cannot identify with certainty the set of all the genuine possible worlds with the collection of all logically possible worlds, even if this has been done quite often in the literature.

However, as long as we consider that the collection of all and only genuine possible worlds includes the actual world, then this collection, independently of whichever M-collection identifies, is said to be alethic. Alethic modality then can be defined in terms of possible worlds as follow:

> "Corresponding to each world and to a set of constraints M, we have the set of M-possible worlds that conforms to those constraints by having all the relevant non-modal statements hold true at them. The alethic modalities are those kinds of modalities for which the actual world is always one of the M-possible worlds that is generates: the constraints generated from the actual world are always satisfied by the actual world" (Divers, 2002, p. 6)

This essentially means that for all alethic modalities the actual world is part of the collection of possible worlds selected by the constraints, so that the actual world is a possible world itself. From this we can deduce that, within alethic modalities, necessity implies actuality and actuality implies possibility: what is true in all possible worlds is true also in the actual world (because this is one of all the possible worlds) and what is true in the actual world is true in at least one possible world (being the actual world one possible world). Therefore, what is necessary is also actual and what is actual is also possible, which corresponds to the ideas presented above that if something is necessary, it is true and if something is true, it is possible.

As we said, the interest of modal discourse is overall on alethic modalities, which are traditionally identified with logical, analytic, metaphysical, and physical modalities, because the actual world is one of the logically, analytically, metaphysically, and physically possible worlds. Non-alethic modalities are instead identified with doxastic and deontic modalities because the actual world does not correspond to the possible worlds generated by our beliefs and laws. Epistemic modality is a borderline case because, if we consider that our knowledge of the actual world corresponds perfectly to what is true in the actual world, that is, if the actual world were effectively one of the epistemically possible worlds, then epistemic modality would be part of alethic modalities.

Within this picture of modality offered by Divers, let us focus a bit more on logical modality, epistemic modality, physical modality, and metaphysical modality (which we will assume here that can be grouped with analytical modality).

Logical modality is the modality that defines necessity and possibility through logical definitions. There is a specific branch of logic, called '*modal logic*', which attempts to provide transcription and calculus for modal expressions of necessity and possibility. The development of the various systems of modal logic has been crucial for the popularity of the possible world account of modality, as a modal logic whose semantics is based on possible worlds has proved to be extremely successful (see the work of Kripke (1963, 1965)). However, we will not venture in the formalism of modal logic here, as for our purposes it should be enough listing the following notions:

- The symbol '□' is the necessity operator, so that '□*p*' means 'It is necessary that *p*'
- The symbol '◊' is the possibility operator, so that '◊*p*' means 'It is possible that *p*'
- The two operators are interdefinable:

  $\Box p = \sim\Diamond\sim p$

  $\Diamond p = \sim\Box\sim p$

(with '~' as symbol of negation)[13]

We could call '*logically possible*' a proposition that cannot be disproved by logical means, and so by using the axioms and laws of logic. Therefore, propositions like "Pigs fly", "Socrates is

[13] The interdefinability of the alethic modalities and of the modal operator is at the basis of Lewis' idea of having might/could-counterfactuals interdefinable with would-counterfactuals, as we saw in section 6 of the previous chapter.

immortal", "Humans have four stomachs" despite being all false in the actual world are all logical possibilities, as none of them can be disproved by the laws of logic. If what is logically possible is only limited by the laws of logic, then logical possibility includes a very large number of propositions, much larger than physical and metaphysical possibility, and this is why it has been sometimes considered to include all that is "genuinely" possible, using Diver's terminology. However, there are propositions that are not logically possible: for example, a proposition including a contradiction like $p \wedge \sim p$ ("It is raining and it is not raining") is *logically impossible*, because a contradiction is a violation of the laws of logic. The proposition $\sim(p \wedge \sim p)$ instead is *logically necessary* because its negation is logically impossible (we can see that $\sim(\sim(p \wedge \sim p))$ is equivalent to $p \wedge \sim p$ which we know is logically impossible).

Epistemic modality defines necessity and possibility on the basis of our knowledge about the content of certain propositions.[14] As we introduced in Chapter I, a proposition is *epistemically possible* if it is compatible with what the speaker knows. When someone says "This milk may be off" for example, and this is given an epistemic reading, the speaker communicates that "This milk is off" could be a true proposition for all they know. This is a possibility compatible with their knowledge of the actual world: e.g., they know that the milk has no label with the best-before date, and that it has been out of the fridge for few days. A proposition that is incompatible with what a speaker knows, as "This cat will lay eggs", is epistemically impossible for that speaker, even if it is still logically possible (there is no violation of logical laws strictly speaking[15]). A proposition is epistemically necessary if its negation is not compatible with what a speaker knows, meaning that it is entailed by what a speaker knows. The proposition "This cat will never lay eggs" is epistemically necessary for those speakers who have knowledge of cats' mammal nature, as this proposition is entailed by the known facts that no mammal lays eggs, and that all cats are mammals; but it is not logically necessary (as we

[14] Again, not to be confused with von Wright's category of "epistemic modality": that was about verification and falsification, while this kind of "epistemic modality" is about compatibility to the speaker's knowledge.

[15] A possible violation could be in the shape a contradiction between the definition of mammals as 'non-laying-eggs animals', the definition of cats as mammal, and the proposition "This cat will lay eggs". However, logically speaking, this proposition can be simply expressed by Pa with Px = x lays eggs and a = this cat, which does not contain any contradiction itself and therefore cannot be disproved by appealing to logical laws only.

saw, its negation above is not logically impossible, even if it is incompatible with the speaker's knowledge).

Epistemic modality is a borderline case as we said above, depending on whether we consider that our knowledge of the world corresponds to how the world is. However, epistemic modality is odd also because it is essentially relative to the single speaker or a group of speakers: propositions are epistemically possible or necessary depending on the knowledge of different speakers. A proposition can be epistemically possible for a speaker or a group and known to be true or known to be false for another: in the milk example, another speaker may have tasted the milk and know whether "This milk is off" is true or false.

Epistemic modality relies on our *knowledge* on how the actual world is, but sometimes we realise that necessity and possibility can rely *directly* on how the actual world is, independently of what we know about it. "No cat lays eggs" is epistemically necessary for most speakers, but we feel that is also something more: in the actual world it will never happen that a cat will lay eggs because this is how the world is, and this is necessary independently of what one knows. If someone, for ignorance, does not know that cats, being mammals, are not physically structured to lay eggs, and so they might accept that "This cat will lay eggs" is an epistemic possibility, this does not change the fact that, in the actual world, it is necessary that no cat lays eggs. This is why epistemic modality is kept separate from physical and metaphysical modality and, as we will see, why Vetter rejects the application of her potentiality framework to this kind of modality.

Physical modality defines necessity and possibility on the basis of how the actual world is, in the sense that what is physically possible is limited by the natural and physical laws of the actual world. A proposition is *physically possible* if is allowed by the natural laws of the world, that is, if it is not against these laws. "There is an atom with atomic number 150" is a physical possibility, as it is allowed by the natural laws, even if it is false that such an atom exists. "Pigs fly", "Humans live forever", "This plane goes faster than light speed" are not physical possibilities, because they are against the laws of nature: pigs do not have the physical features of flying animals, humans physically decay and die, and following the current laws of physics, nothing with a mass can travel faster than light speed. All these propositions are *physically impossible*, even if they are still logically possible (even epistemically possible for some speaker, depending on their knowledge). A proposition is *physically necessary* if its negation is against the laws of nature, meaning that its truth is compelled by the physical laws of the world. "This

rocket will not go faster than light speed" is a physical necessity, because, by nature, a rocket has a mass, and nothing with a mass can go faster than light speed, and this cannot be different in the actual world.

Physical modality seems to account for many of our intuitions around 'necessary' and 'possible': it is less "abstract" than logical modality and less relative than epistemic modality. The relativity here is not anymore on the knowledge of the speaker, but on the world itself: something is necessary or possible because of how things are in the world we live. However, this relativity means that there could be other worlds where things are different and so where other things are necessary or possible. Propositions like, "There could have been flying pigs" or "Immortal humans could exist" are physically impossible, but we would not say that they express something "absolutely impossible". Indeed, as we know, they are metaphysically possible.

Metaphysical modality concedes that the laws of nature could have been different from what they are, so it allows objects to have different properties than the ones dictated by the laws of nature.

Based on what said just above, we could be tempted to say that a proposition is *metaphysically possible* if there could be an alternative version of the actual world in which it is the case[16]. Again though, the best way to explain metaphysical possibility is through examples and comparisons. A good comparison to explain metaphysical possibility is with epistemic possibility. The proposition "The Nazi's concentration camps could have not existed" used by a negationist – a person doubting that the camps existed – expresses an epistemic possibility. On the contrary, the same proposition used by a person who knows and accepts that the camps existed expresses a metaphysical possibility, because they are considering a version of the actual world in which the proposition "The Nazi's concentration camps did not exist" is true, even if it is false in the actual world.

If we accept this conception of metaphysical possibility then a proposition like "These two objects are exactly identical atom-to-atom, but they have a different mass" is *metaphysically*

[16] This does not imply that something that is the case in the actual world is not metaphysically possible. On the contrary, whatever is true in the actual world is a metaphysical possibility, because whatever is physically possible is also metaphysically possible.

impossible (even if is still logically possible) because, even changing the laws of physics, two atomically identical objects cannot have different intrinsic properties, so there is no version of the actual world in which it is true. Similarly, a proposition is *metaphysically necessary* if there is not an alternative version of the actual world in which it is false. "Water is H_2O" is a metaphysical necessity, as there is no alternative version of our world in which it is false.

However, as we know, it is not easy to define exactly what metaphysical modality is, and the definitions we gave above in terms of alternative versions of the actual world could be considered controversial. For example, Divers defines metaphysical modality as the modality concerning "the natures and identity conditions of things" (Divers, 2002, p. 4). Also, he distinguishes between metaphysical and analytical modality, while others think that what Divers calls 'analytical modality' could be included within metaphysical modality. I tend to agree to this idea, because it seems reasonable to think that the case of a married bachelor is in many ways equivalent to a glass of water that does not contain H_2O. Again, this can be controversial, but in this thesis, I prefer to extend the scope of metaphysical modality in this sense.

As we know, metaphysical possibility covers a wider range of propositions compared to physical possibility, but this range is narrower compared to logical possibility. There are more metaphysically possible propositions than physically possible because they can include those going against the laws of nature, but there are more logically possible propositions than metaphysically possible, because there are propositions that do not violate the laws of logic, and so are logically possible, that are metaphysically impossible. "John's drinking water but not H_2O", "That missile blew up Hesperus but not Phosphorus[17]" are all metaphysically impossible, but not logically impossible.

Particularly interesting for counterfactuals are those metaphysical possibilities that are not also physical possibilities, because these are the ones that are against the facts of the actual world, and so they tend very often to be expressed in counterfactual terms. Also, these are the possibilities that cannot be easily explained by the facts of the actual world, and therefore require an alternative explanation.

[17] With Hesperus and Phosphorus being two alternative names of the planet Venus.

This is the reason why metaphysical modality tends to be the focus of modal discourse and why it is the trickiest kind of modality to define and explain. It is because there are metaphysical possibilities exceeding what is physically possible that modality inevitably becomes object of discussion in metaphysics, since it is difficult to explain these possibilities plainly in terms of the actual world. The development of different accounts of modal metaphysics comes from this very issue.

2.2. Modality and Roots

The kinds of modality we discussed above were clearly appealing to a possible world picture of modality, classified depending on the kinds of world over which the quantification happens.

Barbara Vetter (2015) adopts a different type of classification, based on the distinction between 'roots' and 'epistemics'. 'Root modality' is taken to include circumstantial/dynamic modality and deontic modality and is put in opposition to epistemic modality.

The idea is that all modal expressions are used to express different "flavours" of modality, divided into epistemic, deontic, and dynamic:

> "epistemic modality is about what is compatible (or not) with our knowledge, deontic modality is about permission and obligation, and dynamic or circumstantial modality is about developments that are open (or not) given how things really are." (Vetter, 2015, p. 216).

We can explain these three cases with three examples (*Ibid*.):

1. A detective looking at the evidence he has collected in a murder case and saying "Mary might be the murderer" is expressing the epistemic possibility that Mary is the murderer, because it is compatible with this evidence.
2. A father telling his son, who has finished his homework, "You may go out and play now" is expressing a deontic possibility, meaning that the son is permitted to go out and play.
3. A botanist analysing the soil in a foreign country and saying "Hydrangeas can grow on this soil" is expressing a dynamic possibility.

As we will see, Vetter's interest is especially on dynamic modality, which seems to be the closest relative of potentiality because, just like potentiality, dynamic modality is "a matter

simply of how things really are, not how they ought to be or how we know them to be." (2015, p. 216).

The distinction between roots and epistemics is meant to point out the essential difference that there is between epistemic modality and the other two kinds, which are grouped under the class of root modality. This difference is that, while epistemic modality modifies a sentence and deals with the truth value of that sentence, root modality relates the subject of a sentence to an activity and deals with permission, obligation, and ability. Epistemic modality applies to whole sentences and ascribes to them a certain epistemic status of the speaker in relation to them. Root modality, instead, does not take a sentence as a whole to produce a new sentence, but it looks into the structure of the sentence and relates its subject and predicate to each other (see Vetter, 2015, p. 233).

This classification is essentially different from the one above because it contrasts epistemic modality with roots, and then deontic modality with dynamic modality within roots. Rather than focusing on what makes a proposition possibly or necessarily true, this classification is based on the difference of scope over a sentence or parts of a sentence when using modal terms. In this sense, epistemic modality is thought to take a wide scope over whole sentences, while root modality is thought to take a narrow scope over subject and predicate of a sentence. We will discuss this distinction again, but for a more detailed explanation of roots and epistemics see Declerck, (2011).

Nevertheless, the traditional classification in terms of truth may still relate with this one in terms of scope. Vetter's suggestion is that dynamic modality is

> "the closest relative of metaphysical modality, if the latter is understood in a realist way (indeed, some linguists include metaphysical or 'alethic' modality, as they sometimes call it, within dynamic modality)" (Ibid.).[18]

We will come back to dynamic modality when talking about Vetter's account. However, we can already see that she intends dynamic modality as strongly connected to metaphysical modality. Her objective is to offer a new background for metaphysical modality because, as we know, this is the kind of modality most in need of a metaphysical explanation, since it cannot

[18] It will become clear what she means by "realist way" when we discuss Vetter's position on modality.

be so easily reduced to the facts of the actual world. Her idea will be to account for dynamic modality, and so for metaphysical modality, in terms of potentiality.

3. The Problems of Modality

Now that we have considered some of the ways to define modality, we need to discuss some of the questions that modal discourse raises. Indeed, when we talk about modality, we cannot avoid considering certain aspects of modal discourse that touch different philosophical issues, so that modality is a matter of discussion in several branches of philosophy: logic, semantics, epistemology, and metaphysics.

Considering that possibility and necessity are intrinsically connected to truth and have a whole part of logic dedicated to their formalization, modal discourse will of course involve a discussion in logic and semantics. Also, it is natural to ask how we form our knowledge on what is possible and what is necessary, which are epistemological questions. Finally, if we need to account of metaphysical modality, we need inevitably to investigate the nature of the actual world and its entities, whether this world has alternatives and what could be their nature, which are all investigations concerning the metaphysical enquiry. Using Borghini's (2016) terminology, we realise that the Problem of Possibility takes three different shapes:

- The Semantic Problem of Possibility: What does it mean to say that a certain situation is possible?
- The Epistemic Problem of Possibility: How do we come to know that which is possible?
- The Metaphysical Problem of Possibility: What is 'the possible'?

(Borghini, 2016, p. 3)

This clearly shows that 'modality' is nowadays a name indicating a complex and wide series of problems and questions relating to the concept of 'possible'. For our purposes we do not need to explore this complexity in its entirety, but I wanted to make clear that, in this thesis, when talking about 'modality', I will intend it in its contemporary sense, as a category including mainly necessity and possibility, whose relevance ranges from semantics and logic to epistemology and metaphysics.

In this work, while the Epistemological Problem of Possibility will be put on a side, the Semantic Problem of Possibility will be only considered insomuch as it is a necessary step to

develop an account of counterfactuals: in general, once a semantic account of possibility is developed, a semantic account of counterfactuals can be developed on this basis (or at least this is how Vetter puts forward her account of counterfactuals).

The Metaphysical Problem of Possibility consists, essentially, in explaining metaphysical modality, and so in developing a metaphysical background for possibility, both of which are the crucial aims of modal metaphysics. Therefore, in this thesis, we will focus primarily on this problem between the three, because it is the answer to this question that will determine what the metaphysical background of counterfactuals is. If we accept that counterfactuals are part of modality, then a metaphysical account of possibility must extend to counterfactuals as well, so that part of our discussion will focus on whether the metaphysical background we chose for possibility can really apply to counterfactuals too. Given that our choice is Vetter's potentiality account of modality we need to see if this account can successfully apply to counterfactuals.

In discussing modal metaphysics, we will see that accounting for metaphysical modality is not an easy task. Providing a metaphysical background for something that exceeds the actual world can be extremely difficult. The possible world account of modality is appealing for many and has a lot of supporters because it offers a very intuitive picture of how something can go beyond the actual world: by simply being the case in other worlds. However, there are reasons to find it unsatisfactory, especially at a metaphysical level, so that I believe that it is important to explore how an alternative background could work.

Nevertheless, my interest is mainly on how an account of modality relates to an account of counterfactuals. Therefore, it is time to explore a bit more the relation between modality and counterfactuals.

4. Counterfactuals and Modality

Now that we have established what modality is, we should see how this relates to counterfactuals. This thesis supports the generally shared position that counterfactuals are part of modality, given that they require an account of metaphysical modality. Therefore, now we could have a look at how what we learnt on modality can help our understanding of counterfactuals.

In this section we will consider three ways in which modal discourse can influence the discourse on counterfactuals: how the distinction between *de re* and *de dicto* modality extends

to counterfactuals; how the problems concerning modality become problems concerning counterfactuals; and how the appeal to possible worlds can provide an account of counterfactuals.

4.1. *De re* and *De dicto*

Earlier in this chapter we introduced the distinction between modality *de re* and *de dicto* to give account of the fact that necessity and possibility can have two metaphysical readings, one concerning how an object has a property and the other concerning the status of a proposition. Corresponding readings can be found within counterfactuals, as there are both *de re* counterfactuals and *de dicto* counterfactuals. *De re* counterfactuals have the shape "If it were the case that *x* is *F*, then it would be the case that *x* is *G*" like "If John were a bachelor, he would be happier", or "If my cat were not spayed, she would have had kittens with the cat next door". They counterfactually ascribe a property to an object, evaluating the consequences of this ascription. By contrast, *de dicto* counterfactuals have the shape "If it were the case that the proposition *p* is true/false, then it would be the case that proposition *q* is true/false" like "If it were true that someone in this room is a bachelor, it would be true that in this room there is someone who is not married", or "If it were true that unicorns exist, it would also be true that there are horned horses". In this case, they do not ascribe properties to specific objects, rather investigate the consequences of the counterfactual truth or falsity of certain propositions.

An account of counterfactuals that wants to be complete must offer truth conditions for both these kinds of counterfactuals, and we will see that this could create some problems for Vetter's potentiality account.

Note that, in terms of the purely logical reading of the *de re*/*de dicto* distinction, which concerns the scope of the modal operator in relation to the quantifier, counterfactuals could represent a tricky case, since the counterfactual structure can be considered a kind of two-place modal operator. In any case, in general, this thesis will focus on the metaphysical reading of *de re* and *de dicto* counterfactuals, which will be matter of discussion in Chapter V.

4.2. A Problem across Philosophy

We explained above that modality must face the various Problems of Possibility. If we accept that counterfactuals are part of modality, this quite clearly implies that the discussion of

counterfactuals encompasses the same variety of issues that modality does. Therefore, also counterfactuals represent a problem in semantics as much as in epistemology and metaphysics. Indeed, we can reformulate Borghini's Problems of Possibility in terms of counterfactuals. The Problem of Counterfactuals then would take these three forms:

- The Semantic Problem of Counterfactuals: What do counterfactuals mean? What are their truth conditions?
- The Epistemic Problem of Counterfactuals: How does our knowledge of the world influence our beliefs in a counterfactual?
- The Metaphysical Problem of Counterfactuals: What sort of entities should be the ground of the truth of counterfactuals?

Our focus will be on the Semantic and the Metaphysical Problems, leaving out again the Epistemic. Semantically, we will look for a truth-conditional account of counterfactuals. Metaphysically, we will try to understand what the ground of these truth conditions is. Therefore, when we will discuss the potentiality account of counterfactuals, on one side, we will need to discuss the truth conditions that it offers for counterfactuals, and on the other, we will need to investigate the metaphysics of potentiality in relation to these truth conditions.

It is important to say that within the Semantic Problem of Counterfactuals there is also what could be called the 'Logical Problem of Counterfactuals', as the semantic of counterfactuals is accompanied by its own logic. In this thesis though, I will not attempt to develop a new logic of counterfactuals, rather I will explore a particular semantic and metaphysical framework for counterfactuals, leaving to others the problem of seeing if the existing logic could work for it or of developing a brand-new logic if it does not.

4.3. Counterfactuals and Possible Worlds

The fact that counterfactuals sometimes involve metaphysical possibilities that exceed the actual world is one of the motivations that pushed towards a possible world account of counterfactuals. Indeed, one of the options within modal metaphysics for a counterfactual account is to account for the truth of such counterfactuals directly by appealing to alternative versions of the actual world, which are simply other possible worlds.

We have seen above how possible worlds can be extremely useful in defining both the cases and the kinds of modality. The same happens for counterfactuals, as possible worlds have been employed to develop an account of them.

The possible world semantics for counterfactuals is the result of the independent work of Robert Stalnaker and David Lewis.[19] In this section we will first focus on Stalnaker's account, as presented in his paper "A Theory of Conditionals" (1968), and then on Lewis' account as described in his book *Counterfactuals* (1973a). I do not wish to undertake a comprehensive presentation and assessment of them, but since they are the most influential accounts of counterfactuals, I need to address them at least in some details.

Stalnaker's starting point is to consider how we intuitively evaluate a counterfactual and then develop some truth conditions on this basis. He suggests that such an evaluation goes as follows:

> "First, add the antecedent (hypothetically) to your stock of beliefs; second, make whatever adjustments are required to maintain consistency (without modifying the hypothetical belief in the antecedent): finally consider whether or not the consequent is then true" (Stalnaker, 1968, p. 102)

How do we go from this to the truth conditions of a counterfactual? The first step is to find an ontological analogue of our stock of knowledge and beliefs, and this is, in Stalnaker's opinion, a possible world. This is because someone's stock of beliefs can be considered as a way in which this someone represents how things actually are, so when one introduces a false sentence in their stock with the necessary adjustments, it can be considered as a way in which they represent how things could have been. If the original stock can be associated to the actual world, then the modified stock can be associated to a possible world.

Following this, the truth conditions of a counterfactual can be presented as follows:

> "Consider a possible world in which A is true, and which otherwise differs minimally from the actual world. "If A, then B" is true (false) just in case B is true (false) in that possible world." (Ibid.)

[19] The following presentations of Stalnaker's and Lewis' accounts have been inspired by Morato (2019).

An account of this sort, as Stalnaker points out, "has the advantage of providing a ready-made apparatus on which to build a formal semantical theory" (1968, p. 103), constituted by modal logic and its possible world semantics. His proposal is to extend the semantics of modal logic by adding a selection function *f*, which takes a proposition and a possible world as arguments and a possible world as value. The expression *f*(A,*w*) should be read as "the selected world possible relative to the base world *w* in which A is true". This function selects, for any antecedent A, a possible world in which A is true that is the *most similar* to the base world *w*. A counterfactual "If it were that A, then it would be that B" is true when the consequent B is true in the world selected by the function. Adopting '>' as the counterfactual corner, we can state the following semantic rules:

A > B is true in *w* if B is true in *f*(A,*w*)

A > B is false in *w* if B is false in *f*(A,*w*)

Following this interpretation, Stalnaker can show how the logic of counterfactuals is an extension of modal logic. If traditional modal logic allows us to talk about what is true in the actual world, in all possible worlds or in some unspecified world, the selection function added to the semantics, together with the counterfactual corner added to the logical language, allows us to talk about what is true in specific possible worlds.

We will not discuss his account in more details, but we need to fix some very important points of Stalnaker's theory: first, that he considers that there is only *one world* that is the most similar to *w* in which A is true and, second, that the selection function creates a total ordering of all possible worlds relatively to *w* based on *similarity*, with *w* as first element.

Lewis developed his semantics independently of Stalnaker, but we can use a presentation in terms of selection function just the same, to make the comparison easier. Lewis' preferred exposition is much more complex in a way, using the so-called 'sphere' representation for indicating various degrees of similarity between worlds, which he calls 'closeness'. For our purposes, we will try to keep it simple focusing on the crucial differences between the two semantics.

It is important to point out that, while Stalnaker was interested to an account of conditionals in general, and mostly talking in those terms, Lewis' focus is on counterfactuals only. This reflects also in the formal language adopted by them. Stalnaker introduces the 'conditional' corner '>' (we called it 'counterfactual' corner earlier only for simplicity) aiming

to use it as operator for all conditionals. Lewis, on the contrary, introduces two 'counterfactual conditional operators', corresponding, as we know, one to would-counterfactuals and the other to might-counterfactuals (Lewis, 1973a, pp. 1-2):

"A$\Box\!\!\rightarrow$B" must be read as "If it were the case that A, then it would be the case that B"

"A$\Diamond\!\!\rightarrow$B" must be read as "If it were the case that A then it might be the case that B"

Recalling what said in Chapter I about might-counterfactuals (section 6), Stalnaker denies that there is a real difference between them, relegating might-counterfactuals to mere expressions of the speaker's epistemic status towards the counterfactual. Lewis instead, not only links to each a different logical operator, but also offers different truth conditions for them. He also shows how the two operators are interdefinable, as we know:

$A \Box\!\!\rightarrow B =_{df} \sim(A \Diamond\!\!\rightarrow \sim B)$

$A \Diamond\!\!\rightarrow B =_{df} \sim(A \Box\!\!\rightarrow \sim B)$

Even if we have already introduced them in Chapter I, here we can explain the different truth conditions of would- and might-counterfactuals more in depth. To do so, we need to point out one of the crucial differences between the systems of Stalnaker and Lewis. In Lewis' semantics there is no condition imposing that there can only be *one world* selected by the function in which A is true that is the most similar (or "closest") to *w*. Therefore, for Lewis, we can select multiple possible worlds for the antecedent of a counterfactual, so that we can identify a *set of worlds* in which A is true that are the most similar to *w*.

Following this, we can define the truth conditions that can be found in Chapter I:

I. A counterfactual 'A$\Box\!\!\rightarrow$B' is true at world *w* iff:
   i. there is no possible world in which A is true (trivially true);
   ii. or, all the possible worlds closest (most similar) to *w* in which A is true are worlds in which B is true.

II. A counterfactual 'A$\Diamond\!\!\rightarrow$B' is true at world *w* iff there are some possible worlds sufficiently close (similar) to *w* in which A is true that are worlds in which B is true.

We can see that for Lewis, contrary to Stalnaker, there is a quantification over the worlds in which the antecedent is true and that are the most similar to *w*, and it is this quantification that makes true either would- or might-counterfactuals.

With this short presentation, we have introduced what we call the 'Possible World Account of Counterfactuals' (PWA) using this term as a label including both Lewis' and Stalnaker's accounts. Naturally, this is a very reductive introduction, but it should show the main ideas behind PWA, and why it became so popular: on one side, the notion of possible world has a strong intuitive foundation concerning counterfactuals, as our counterfactual thought process mostly involves the formulation of hypothesis which presupposes our capacity of imagining alternative ways in which the world could be; on the other side, both the semantics and the logic of PWA can be presented as extensions respectively of the more general possible world semantics and of modal logic, so that they can be included in a well-established theoretical framework.[20]

In this work, I will not try to disprove PWA, because it is not my objective. The aim of this thesis is to discuss how well an account of modality based on potentiality can apply to counterfactuals, so I do not need to discuss possible worlds. However, I can try briefly to motivate my choice of abandoning an account of counterfactuals that is generally considered quite appealing.

My idea is that PWA can be criticised in two ways: either we criticise the semantic treatment of counterfactuals within the account, or we criticise the initial premise of the account that possible worlds should be implemented as its metaphysical basis. In this sense, the aim of this thesis is not to criticise the semantics itself, but to investigate an alternative metaphysical background, which in turn brings along a different semantics. Therefore, here I am not interested in disproving the possible world semantics of counterfactuals. Indeed, I could potentially agree that it is an effective semantics within the possible world perspective. What I consider to be the problem with it is its initial assumption that we need possible worlds.

[20] I owe most of my knowledge and understanding of the basics of possible world semantics to the introduction offered by Paolo Casalegno (1998). For a good literature on all the issues raised here you can refer to: Wittgenstein (1922), Carnap (1947) and Negri (2009) for possible world semantics and modal logic; Lewis, & Langford (1934), Ramsey (1931), Chisholm (1946) and Goodman (1947) for accounts of counterfactuals before possible worlds, Fine (1975) and Kratzer (1989) for some criticism and alternatives to PWA.

In the same way as I do not want to disprove PWA, so I will not linger too much on justifying my rejection of possible worlds as metaphysical background of counterfactuals. In a way, I am simply aligning myself to Vetter's own position, as we will see in the next chapter.

Very simply, I support the position taken by the so-called 'new actualists' (see Vetter, 2011), who suggest that modality is better explained in terms other than possible worlds, like essences or potentials. The reason for this is that the metaphysical definition of possible worlds brings along some ontological costs that for some, like Vetter and me, are unacceptable. In discussing Vetter's account, we will analyse her motivations for preferring not to use possible world, but in general, I agree with her in preferring not to have to talk about possible worlds when talking about a background for counterfactuals.

However, I consider myself a sort of 'selective-abstentionist', meaning that I would like to abstain from the use of possible worlds mostly in relation to a counterfactual account. Of course, following and supporting the position taken by new actualism, I will suggest that the best way to do this is to find an alternative metaphysical explanation of modality altogether, which essentially removes possible worlds from their most important role. However, I do not want to say that possible worlds cannot help in other areas or at least in a heuristic way that simplifies and makes clear certain explanations. I simply reject their role in the discussion about counterfactuals, supporting the idea that a different metaphysical background is worth investigating.

My aim is not to deny the possible world doctrine altogether, rather is to see whether an alternative view on modality not involving possible worlds, but potentiality, would serve the counterfactual discourse more efficiently in ontological terms. Note that some new actualists too do not want to proceed in a crusade against possible worlds, rather their main objective is to consider alternative ways to account for modality. New actualists do not want to justify their position by disproving the possible world discourse, rather by proving that there is an alternative. This is what this thesis will try to explore concerning counterfactuals.

With this, we can conclude our discussion of possible worlds and counterfactuals: there is a possible world account of counterfactuals, and it works quite well, but in this work, we wish to focus on another account, to see if it can work even better.

5. Conclusion of Chapter II

In this chapter we have investigated the relation between modality and counterfactuals. After introducing and clarifying many aspects of modality in sections 1 and 2, along with showing the problems that modality raises in section 3, in section 4 we discussed how the fact that counterfactuals are included within modality can influence the discourse on counterfactuals.

The inclusion of counterfactuals within modal discourse was motivated in the last chapter by the fact that counterfactuals sometimes involve metaphysical possibilities, and therefore require an account of metaphysical modality for the formulation of their truth conditions.

However, there are other motivations that encourage a close relationship between counterfactuals and modality. For example, as we explained in the last chapter again, it seems that counterfactuals are not comparable with other conditionals and cannot be satisfactorily explained in those terms, so another comparison is needed. We can consider that counterfactuals have two crucial features: they are a kind of *conditional* and they require a discussion of *possibility* to be assessed. If we appeal to the first feature, then we should try aligning them with the discussion of indicative conditionals. However, if this does not work for producing a satisfactory account, then the next natural choice is to appeal to the second feature and include them in the discussion over possibility. Because possibility is part of modality, and the indicative conditional framework do not seem to offer a satisfactory account, modality seems to be the best candidate to offer a good account for counterfactuals.

Instrumental to the fact that counterfactuals can be considered part of modality is also the idea that modality as a category does not include only necessity and possibility. As a matter of fact, even if modality has become more and more centred on necessity and possibility, at the same time it has been suggested that these two are not the only components of modality, as there are other entities which are part of this category. It is thought that entities like essences and dispositions are part of modality as well as necessity and possibility[21]. This extension of the modal 'package', as Barbara Vetter (2011, 2015) likes to call it, also includes counterfactuals. In addition to this, it is crucial to point out that, like necessity and possibility, also these "other

[21] I will not explain what essences and dispositions are here, as they will be discussed in depth later in the thesis.

modalities" are not considered separated and independent of one another; on the contrary, many want to extend the idea of interdefinability also to these less familiar entities. For example, counterfactuals have been used to explain dispositions, essences have been explained through necessity, and so on. By including counterfactuals in this "modal family", we might find that we can explain them through other members of this family, which, as we will see, is the idea at the core of the potentiality account of counterfactuals.

Part of the next chapter will be on this last point. We will see how a different conception of modality as a larger category including dispositions and essences gives a push towards new accounts of modality. This shift is at the origin of the potentiality account of counterfactuals, as we are going to see.

CHAPTER III

COUNTERFACTUALS AND DISPOSITIONS

In the last two chapters, we have seen that counterfactuals can be included in modal discourse because they seem to require an account of metaphysical modality. After making clear what counterfactuals are, and what their relationship with modality is, it is time to present the account of metaphysical modality that we will support as framework for counterfactuals. This account is, as we know, Barbara Vetter's version of dispositionalism, which can be called 'potentialism'.

The aim of this chapter and the next is to introduce and discuss Vetter's potentiality account. In doing so, we will focus on her work *Potentiality. From Dispositions to Modality* (2015), which includes her initial proposal of an account of counterfactuals in terms of potentiality. The presentation of the Vetter's material will help us understanding her metaphysical and semantic framework so that we can make an informed evaluation of her proposal concerning counterfactuals. However, we will not focus too much on Vetter's account of modality in general or on the possible general criticisms against it, as our interest is mainly on applying this account to counterfactuals.

This chapter intends to be preparatory for a more detailed discussion of potentiality in relation to counterfactuals happening in the next chapter. Therefore, here we will introduce dispositionalism in general and explain why Vetter preferred moving towards potentialism. Also, we will explain better the motivations behind her rejection of a possible world account of modality, which will reflect our own. Finally, in this chapter, we will simply outline Vetter's proposal concerning counterfactuals, with the clarification and the justification of some metaphysical and semantic aspects of this proposal being discussed in the next.

Dispositionalism is, as the name suggests, a group of theories trying to explain dispositions, as well as trying to explain other things through dispositions. There are various positions that one can take on dispositions, and Vetter is very clear on her place in the debate: she is a non-reductionist and realist about dispositions. Her aim is to generalise the notion of disposition towards the notion of potentiality, which, in her view, represents the metaphysical background not only of dispositions, but of possibility in general, in substitution of possible worlds.

The structure of the chapter will be the following. In the first section we will introduce dispositionalism in general as a form of new actualism and Vetter's position within this philosophical trend. The second section will present some of the motivations behind our rejection of possible worlds, and why we think dispositionalism is preferrable. In the third section, we will explain how Vetter's dispositionalism is different from others. In the fourth section, we will finally see what her account of counterfactuals is. The final section will conclude the chapter in preparation for the next.

1. Introduction to Dispositionalism

The discussion of modality, as we know, raises some very important questions, specifically in relation to metaphysical modality. A way to answer these questions is in the shape of operators defined in terms of possible world, as the possible world account of modality suggests. Another possible way to find these answers is to focus on a particular kind of modal notion, that is called 'disposition'. The term 'dispositionalism' is used to refer to the group of philosophical theories that focus on dispositions.

However, with 'dispositionalism' we can intend two very different approaches to dispositions which must be clearly distinguished. To understand these two approaches, we must first clarify two ways in which philosophers have tried to explain modality: by reducing modality to the non-modal and by reducing modality to the actual. Philosophers trying to reduce modality to the *non-modal* want to exclude any modal notion from their explanation of modality so that they deny that modality should be explained in terms that are themselves modal. The main supporter of this view is David Lewis, whose appeal to a quantification over non-actual concrete worlds, called 'modal realism', aims to exclude any modal notion from the account of modality. Philosophers trying to reduce modality to the *actual*, instead, do not aim

to exclude modal notions from their explanation, rather they want their account not to include anything that is non-actual, so that modality should be explained entirely in terms of the actual world. This view is generally called 'actualism'.

With this distinction in mind, we can explain the two approaches to dispositionalism. On one side, there is a dispositionalism that is a form of actualism, whose aim is to reduce the modal to the actual via the appeal to dispositions and which is supported by Vetter. On the other side, instead, there is a dispositionalism that appeals to possible worlds, wanting to reduce dispositions to the non-modal rather than to the actual via a form of quantification over worlds. While the former tries to present an alternative way to explain modality through dispositions, the latter's main aim is not to find an alternative account of modality in terms of dispositions, rather to explain dispositions by applying the possible world apparatus to them.

The main difference between these versions of dispositionalism can be explained also in terms of metaphysical *reductionism* and *realism* about dispositions: reductionism considers dispositions a sort of construction that need reducing to other modal notions (e.g., possible worlds and counterfactuals), while realism embraces the opposite idea of accepting dispositions as primitive real entities.

Vetter identifies herself as realist dispositionalist who pursues a form of actualism. In particular, she calls the kind of actualism that she supports 'new'[22] actualism to contrast it with the traditional 'classical' actualism (2011).

In explaining 'actualism' in general, Vetter suggests that it is intended primarily in opposition to Lewis' modal realism and to Quine's scepticism. In this sense, the general approach of actualism is to take modality to be real (against scepticism) and consider that modality can be explained by the content of the actual world only (against modal realism). With this purpose, 'classical' actualism tries to reduce the modal to the actual by providing an account of possible worlds wholly based on the actual world. What Vetter calls 'new' actualism, instead, abandons this idea, thinking that possible worlds, even if they are a useful tool in modal logic, "they have little to do with the metaphysics of modality" (2011, p. 742). Therefore, new

[22] Note that her use of 'new' must not be intended as a novelty in a temporal sense, since it has actually been around for quite a while (see e.g., Forbes, 1985 and 1988). Rather, this is a label devised only to create a contrast with the more classical possible-world-centred version of actualism.

actualism looks for the grounds of modal truth within the actual world, but without appealing to possible worlds, so that it appeals to other modal notions. The general idea of new actualism is that the main kinds of modality – necessity, possibility, and counterfactuals – are just part of a bigger "package", as Vetter calls it (2011, p. 743 and 2015, p. 4), constituted by essences and dispositions as well, and that the main kinds can be accounted for by these other parts of the package. Therefore, the ideal aim is to explain necessity and possibility by using essences and dispositions.

Vetter's preferred version of new actualism commits to a view of dispositions and essences as fully part of actuality, being properties of actual concrete objects. Because, within this view, dispositions and essences are always 'of' something, in the sense that they are always related to the individual objects of the actual world, the consequent thought is that if modality is grounded in dispositions and essences, then modality too is grounded in the actual world. The focus on essences or dispositions divided new actualism in two main branches: essentialism and dispositionalism. Only the latter is the object of our discussion, so we will not discuss essentialism here.

Now that we have established that Vetter is a realist about dispositions and a supporter of the idea that modality can be explained by the dispositions of the objects in the actual world, we can try to make sense of what we mean by "dispositions".

What dispositions are and how we should define them is quite controversial. Tim Crane (1996) gives this definition: "a disposition is a property (such as solubility, fragility, elasticity) whose instantiation entails that the thing which has the property would change, or bring about some change, under certain conditions" (p.1). When an object is said to have a disposition, it generally means that, under certain circumstances – i.e., in the presence of a certain 'stimulus' – the object will behave in a particular way – i.e., a certain 'manifestation' will follow (see Wang, 2015, p. 454). Crucially, an object can possess many dispositions without manifesting them: a glass has the disposition of fragility even if is not broken. In this sense, the manifestations of dispositions can be *merely possible*. It is this connection between dispositions and possibility that brought dispositionalists to investigate whether an account of possibility could be formulated in terms of dispositions, as we will see.

Another way to define dispositions is through the conditional analysis, as explained by Vetter in (2011, p. 748) and then extensively in her book (2015): given that *M* is a certain manifestation and *S* is a certain stimulus, *x* is disposed to *M*, iff, if *x* were *S*, x would *M*. For

example, fragility, the disposition to break, is expressed by the counterfactual conditional "If *x* were struck, *x* would break". Since the conditional analysis is based on the association of dispositions with some counterfactuals, this initial approach to dispositions was a reductive one, trying to reduce dispositions to these counterfactuals.

However, the conditional analysis was challenged by appealing to "finks" and "masks" (see Martin, 1994): a fragile glass may be struck and not break because of some external condition, so that the fragility of the glass can be "masked", for example, if it is packed in Styrofoam. For the same reason, the reduction of dispositions to counterfactuals fails: a counterfactual like "If this glass were struck, it would break" may be false for many external reasons, like if again the glass is packed in Styrofoam, yet the glass would maintain its disposition of fragility.

Therefore, some dispositionalists changed direction towards a non-reductive account, suggesting that counterfactuals are not the base of dispositions, rather that it is dispositions that make counterfactuals true (see Bird, 2007 or Jacobs, 2010). In this case, the idea that a disposition is generally characterised by a counterfactual is maintained, but instead of having dispositions reduced to counterfactuals, as the conditional analysis suggests, the order is reversed, so that a dispositionalist theory of counterfactuals can be formulated as follows:

> (C) "A counterfactual of the form 'If *x* were *S*, then *x* would be *M*' is true just in case *x* has a disposition to *M* if *S*". (Vetter, 2015, p. 13)

Within this view, the truth or falsity of a counterfactual is explained by *x* having or lacking a certain disposition. However, even (C) must fail, again because of the problem of masks and finks. In general, as Vetter argues (2015, p. 14), a disposition is a matter of how things stand with a particular object, being an intrinsic property of this object, while the truth of a counterfactual depends on more than the intrinsic nature of a particular object, like some external factors or some extrinsic properties, so that the truth values of a disposition ascription and the corresponding counterfactual can be different. Therefore, both the conditional analysis and (C) are inadequate as they stand.

Nevertheless, Vetter embraces the non-reductive stance concerning counterfactuals that lead to the formulation of (C), so that, in rejecting the conditional analysis, her dispositionalism wants to offer some truth conditions for counterfactuals in terms of dispositions, rather than trying to explain dispositions through counterfactuals. The way she does it, however, is not

based on dispositions directly, but on what she takes to be their metaphysical background, potentiality, as we will see.

With the rejection of the traditional connection between dispositions and counterfactuals, dispositionalists looked for another non-reductive approach to dispositions, still aiming to use dispositions to account for modality. Indeed, even if dispositions seemingly cannot explain (nor be explained by) counterfactuals, this does not exclude that dispositions can be used to explain other modal notions. Therefore, an alternative form of dispositionalism suggests that dispositions should be used to provide an account of *possibility* rather than of counterfactuals. This idea is based on the fact that objects can possess many dispositions without manifesting them, so that the manifestation of these dispositions is *merely possible*. On this basis, dispositionalists could abandon a stimulus-manifestation view of dispositions and suggest that dispositions can be identified only by their manifestation. In this way, an account of metaphysical possibility based on dispositions could take the following shape:

> (P) "A possibility statement of the form 'It is possible that *x* is *M*' is true just in case *x* has a disposition to be *M*". (Vetter, 2015, p. 14)

In this case, in contrast with (C), it is the truth or falsity of a possibility claim that is explained by *x* having or lacking a certain disposition, rather than the truth or falsity of a counterfactual.

This dispositionalist account of possibility is the starting point of Vetter's potentialism. However, before discussing how Vetter's specific form of dispositionalism develops, we could wonder why we should support such a dispositionalist view of modality rather than the more traditional possible world account.

2. Dispositions vs Possible Worlds

We said already that possible worlds have been a popular choice in modal discourse because of their theoretical utility, since they provide both an effective semantics for modal logic, by linking possibility and necessity to existential and universal quantification, and an efficient formal model for explaining different kinds of modality, by introducing restricted quantification.

However, I also said that I agree with Vetter in thinking that possible worlds are not as effective tools in terms of modal metaphysics, and now we can finally investigate a bit more why I support this claim. Indeed, even if the formal model attached to possible worlds is extremely powerful, I do think that its metaphysical background is in some ways inadequate. As Vetter points out, there are "doubts that possible worlds really provide an insight into the nature of metaphysical modality" (2015, p.6).

The source of this metaphysical dissatisfaction can be found in the debate arising from the crucial question concerning the nature of possible worlds, and in two very different answers that can be given to this question: whether they are *concrete objects*, as modal realism claims, or *abstract objects*, as classical actualism maintains.

In what follows, we will discuss some objections against both views. I do not want to explore these criticisms too deeply, as it is a very complex debate and covering it all would take us away from the focus of this work. However, this discussion can at least show that there are reasons to prefer an alternative to the metaphysics of possible worlds. In addition, we will see why Vetter thinks that dispositionalism is preferrable, and which advantages such a theory should have over the possible world framework.

2.1. Against Modal Realism

The most prominent supporter of modal realism (also called "concretism" or "genuine realism") is David Lewis (1973a and 1986). His starting point is to define the actual world as 'the way things are'. It is widely acceptable that things might have been different and that there are infinite alternative ways in which things might have been. Between the various ways in which things might be, there is the way things are, which is, as said, the actual world. If 'the way things are' is a world, then all these 'ways things might have been' are worlds as well. They are what we call 'possible worlds' (see 1986, p. 1).

Lewis' crucial idea deriving from this point is that possible worlds are of the same kind of the actual world. The actual world is in fact only a world between many, and they are all worlds in the same way and with the same *concrete* nature, with the only difference that ours is 'actual'. However, our world is 'actual' only because is the one to which we all belong. Other worlds are 'actual' for their inhabitants: actuality, in this sense, is relative (see 1973a, pp. 85-86). Therefore, if the actual world is "us and all our surroundings" (1986, p. 2), possible worlds

are the same thing for their inhabitants. Possible worlds then exist in the same way our world exists: for Lewis they are like "remote planets" (*Ibid*.).

However, in Lewis' view it is equally crucial that possible worlds are isolated. We would never make contact with another possible world, and there is not any kind of causal relation between any world and any other: something happening in a possible world never influences what happens in another. Also, there is not any causal role between worlds: no world is generated by another.

Because a world cannot generate another, Lewis rejects strongly the idea that possible worlds are our own making (1986, p. 3): they are not products of our mind, nor they can be reduced to something like linguistic entities (1973a, p. 85). Possible worlds are worlds in the true sense of the word, concrete objects, exactly like our own.

These are the main ideas of modal realism, which can be summed up as the "thesis of plurality of worlds (…) which holds that our world is but one world among many" (1986, p. 2): the idea that possible worlds exist exactly like our own world, but they are actual only relatively to those who inhabit them, so that only the actual world is actual for us.

Modal realism is considered a controversial position and quite extreme. The fact that possible worlds should be taken to be *concrete* raises some strong objections. Vetter recognises some of these objections and takes them as good reasons to pursue a metaphysics of modality that is not based on possible worlds. Here we will just introduce four of these objections and we will not discuss the possible answers offered by Lewis.

The first objection that we can consider is the widely recognised "ontological extravagance" of modal realism (see e.g., Divers, 2002, p. 154): such a view could lead to the idea that everything that is possible belongs to the realm of concrete existence, creating an infinite proliferation of other universes and making reality overpopulated. This is due to the vast number of things, the infinite concrete possible worlds and everything in them, that modal realism suggests existing, going far beyond what is commonly believed to belong to the realm of existence.

The second objection to consider, and probably the one that is primarily guiding our rejection of modal realism, is the so-called 'incredulous stare' (named by Lewis himself in 1973a, p. 86 and 1986, p. 133). The incredulous stare is identified with the negative reactions of incredulity and incomprehension caused by modal realism when considering its ontology

(see for example, Divers, 2002, p. 153). It is recognised by Lewis that his ontology is radically against common opinions about reality and that this represents a cost of his theory. Therefore, modal realism could be criticised as being incredible *prima facie*, so that this would be a prohibitive and irrecoverable cost of this view.

The third objection is related to the second: even if we were open to accept the implausible and incredible ontological picture that modal realism puts forward, still we seem unable to find some sound evidence that could support it. The main epistemological criticism against modal realism is indeed that "we cannot know that its characteristic ontological claims are true" (Divers, 2002, p. 149). Given that the concrete possible worlds are completely isolated from each other, and so that other worlds are entirely out of our reach, how can we gain knowledge of them or simply establish that they are indeed there?

The fourth objection that we will consider here is that, even if it were true that there are infinitely many concrete worlds, this does not seem to be a fact about possibility and necessity, but rather a fact about the one actual world, which includes these "universes". Indeed, this fact about our world seems entirely contingent and could be considered uninteresting for a theory of modality, with all these universes not really being relevant for determining the truth of modal claims.

To be fair to Lewis, he battled very hard to answer to all these objections and some of his solutions are quite effective. However, in general, I agree with Vetter that these objections give us enough reasons to prefer investigating a different metaphysical background for modality, and so we will not enter the debate on modal realism here.

2.2. Against Actualist Realism

We have seen the objections that can lead one to be sceptical about possible worlds being concrete. However, possible worlds cannot escape scepticism even if they are taken to be abstract, as classical actualism suggests.

Classical actualism, which has been called 'actualist realism' by Divers, supports the idea that possible worlds exist in the sense that they all actually exist, but only one of them, the actual world, is actualised. In this sense, possible worlds are all abstract objects with only one possessing concreteness.

> "The actualist realist (AR) takes unrestricted quantification to range over all and only the things that actually exist. AR […] affords (at least) four paradigm conceptions of the sense of the 'world' predicate. According to these respective paradigms 'worlds' refers to: (a) a certain kind of state of affairs (Plantingan realism, PR); (b) a certain kind of recombination of the actual individuals and the actually instantiated properties (combinatorial realism, CR); (c) a certain kind of complex property (nature realism, NR); (d) a certain kind of interpreted sentences (book realism, BR)." (Divers, 2002, p. 21)

As we can see, there are various forms of actualist realism, depending on which abstract entities the possible worlds are identified with. We will not discuss each of these forms because the criticism we can move towards actualist realism is quite general, and can apply to all of them, because it is based on the general idea that possible worlds are abstract objects.

For example, Vetter's criticism (based on Lewis 1986, Williamson 1998, and Jubien 2007) is that if we consider possible worlds to be abstract entities, like sets of propositions, then we must find ways to distinguish those sets that do correspond to genuine possibilities to those that do not:

> "If abstract possible worlds are supposed to deliver a robust account of metaphysical modality, it is hard to see how they can avoid circularity; if not, then they are simply irrelevant to the metaphysical question of what possibility and necessity are" (Vetter, 2015, p.6,).

There are other ways to criticise actualist realism, including objections to each form of it, but we will not consider them here. We will just say that the fact that actualist realism considers possible worlds to be abstract objects is not so much better than considering them concrete, because abstract entities raise quite a few metaphysical questions themselves. It has been argued that actualist realism can compete with modal realism in terms of ontological costs, by offering an actualist ontology of possible worlds that seems more acceptable than the non-actualist ontology implied by modal realism. However, my idea is that the ontologies relating to each form of actualist realism, given that are all based on some abstract entities, are not so obviously safer than Lewis' concrete view of possible world. Even if it could be argued that the potential problems concerning abstract objects do not themselves undermine the view that possible worlds are abstract, they at least provide motivation for exploring an alternative approach, Therefore, I maintain that a modal metaphysics not involving possible worlds is worth investigating.

In general, I agree with Vetter in concluding that we have good reasons to doubt that "the formal apparatus of possible worlds can be simply implanted into a metaphysics of modality" (2015, p. 6), both in terms of modal realism and of actualist realism, and this explains my support to the 'hardcore actualism' that is new actualism, which refuses any appeal to possible worlds, whether abstract or concrete.

I do think that the talk about possible worlds can still maintain its place as a formal model or a descriptive and heuristic tool, but I align myself with Vetter in suggesting that it cannot aspire to a place in metaphysical modality. As she says:

> "[…] we should not make the mistake either of thinking that 'possible worlds' are genuinely worlds, or that they have any special connection with possibility. […] Thinking that there is a tight connection between metaphysical modality on the one hand, and possible worlds as a formal tool on the other, may prove unhelpful both in accounting for metaphysical modality and in using the tool of possible worlds in other areas" (Vetter, 2015, p.10).

Now that we have seen some reasons to reject possible worlds, we should have a look at the reasons to support dispositionalism as an alternative.

2.3. In Support of Dispositionalism

Vetter claims (2015, p. 11-12) that a dispositionalist theory of modality has many advantages both in terms of metaphysics and in terms of epistemology. While I agree that its metaphysical virtues are quite compelling, I have some doubts that a dispositionalist epistemology, at least as Vetter puts it, is as easily acceptable and as evidently empiricist as she suggests. However, such an epistemology does look promising compared to a possible world-based one.

In terms of metaphysics, I can see that dispositionalism is, on one side, very parsimonious because if we already have dispositional properties, and we do not need anything else to account for modality, then we do not need to postulate the existence of other entities when explaining modality. On the other side, I also agree that its ontological picture is extremely intuitive, because it is a metaphysical realism (i.e., the idea that the world is constituted by objects and properties), and because it anchors possibilities to objects that are not "otherworldly" but are the ordinary objects of the actual world, so that it can avoid the appeal to entities other than the ones belonging to the actual world when explaining modality.

By being parsimonious in the number of entities it presupposes existing, dispositionalism seem to gain an advantage over modal realism. At the same time though, by assuming that dispositions are concrete – in the sense that they are real properties of the concrete objects of the actual world, whose possession can clearly have concrete consequences on these objects, like the fragility of a glass which can result in its breaking – it seems able to maintain the advantages of appealing to concrete rather than abstract entities, in contrast with actualist realism[23].

In terms of epistemology, Vetter's claim (see 2015, p. 11) is that, by anchoring possibilities to the ordinary concrete objects, dispositionalism seems able to offer a plausible epistemological picture as well. She believes that we do have secure knowledge of dispositions of the individual objects as it is a knowledge coming from our everyday context. She does not really explain where it comes from, but she considers that this knowledge we have of dispositions has an empirical nature (see 2015, p. 12). If, this is the case, then, if metaphysical modality is based on dispositions, then again, our knowledge of these dispositions is a way of knowing about metaphysical modality. And if we already have knowledge of metaphysical modality when we have knowledge of dispositions, then we have in dispositions a direct source of knowledge of metaphysical modality. In this way, Vetter claims that dispositionalism can offer simplicity where the other views cannot. On one side, modal realism clearly struggles to reconcile its metaphysics of modality with its epistemology because it cannot provide a clear connection between the completely inaccessible concrete worlds and our knowledge of them. On the other side, the forms of actualist realism too, by considering the worlds as abstract objects, struggle with their epistemology, because these possible worlds are as well very remote from our everyday modal knowledge.[24]

Vetter's view on the epistemology connected to dispositionalism is in many ways excessively optimistic because we could be sceptical that our knowledge of dispositions is so easily acquirable by empirical means. Even if it is true that we do have empirical knowledge of

[23] Again, abstractness might not be a fault itself, but concreteness seems to make dealing with these entities much easier.

[24] Note that solutions to the epistemological problems of both modal realism and actualist realism have been put forward, but for simplicity we will simply agree with Vetter on this point.

the objects possessing dispositions, still we might wonder if this means that we know those dispositions in the same way.

However, this does not mean that an epistemology of modality based on dispositions is not promising: dispositions, being conceived as so strongly related to concrete actual objects, can surely be good candidates for more knowable entities than Lewis' concrete worlds or actualist realism's abstract worlds. Even if I do not want to go as far as Vetter in enthusiastically supporting this epistemology as the simplest – to the extent of taking it to be an empiricist epistemology – her program does seem to take promising steps towards a possible integration between metaphysical modality and its epistemology.

Bearing this objection in mind, still we can see why Vetter suggests that a view of dispositionalism as an object-property ontology with an empiricist modal epistemology is extremely appealing, because it can offer both an intuitive ontology and a modal epistemology that can potentially solve the integration challenge with its metaphysics. Even if I can still have doubts about its epistemology, I recognise the appeal of a dispositionalist theory of modality in contrast with the traditional possible world account and I agree that such a theory should be pursued to produce a new framework for counterfactuals.

Now that we have explained, at least in broad terms, why I support her metaphysical preferences, we can finally introduce Vetter's own version of dispositionalism.

3. Vetter's Dispositionalism

It is clear that, to understand how Vetter's account of counterfactuals comes about, we must understand her dispositionalism first. Her approach to dispositions is different from others because she develops it by considering the more basic and intuitive idea of 'potentialities', hence potentialism: potentialities include dispositions, like fragility, but also abilities, like playing the piano, and potentials, which are weaker than dispositions (e.g., the potential to break is "weaker" than fragility). Therefore, every disposition is a potentiality, but not every potentiality is a disposition, because some potentialities are abilities, while some others are potentials. In this way, we can see that potentiality is a widespread category, which Vetter suggests includes all those properties that are ascribed by the predicate 'can' in its dynamic reading, so that she identifies a property as a potentiality if it concerns "what a given individual *can* do" (Vetter, 2015, p. 1).

Being a non-reductionist and a realist, she maintains that potentialities should be considered real entities that can serve as a *source of explanation* in metaphysical modality, rather than something which needs an explanation. Vetter adopts the term 'potentiality' because it gives a general idea of what she talks about, but she does not intend to give an explicit definition of this term. Indeed, Vetter's idea is to take the notion of potentiality as *primitive*, being a generalisation of the notion of dispositions, and she wants to explain metaphysical possibility through this notion. Her ultimate aim then is to have an account of modality based entirely on potentiality.

Despite assuming it to be primitive, Vetter tries to explain what potentiality is and her starting point is to compare it with possibility. While potentiality is closely related to possibility, there is a difference between the two that can be compared to the difference between localized and non-localized modality (see Vetter, 2015, p. 2):

- Potentiality is *localized* because it is a property of a particular object (e.g., a glass *has* the potentiality to break).
- Possibility is *non-localized* because 'being possible that *p*' is not necessarily a fact about a particular object but a fact about how things could turn out to be (e.g., it is possible *that* it rains tomorrow).[25]

The idea behind her account then is to explain non-localized modalities, such as possibility and necessity, by using localized modalities, in the shape of potentiality. Potentialities in this sense are intended by Vetter as individuated only by their manifestation, abandoning the traditional picture of dispositions as being characterized by a stimulus and a manifestation. In the same way as the disposition used in (P) above to verify a possibility statement is simply a 'disposition to M', and not a 'disposition to M if S', so the potentialities used by Vetter to explain possibility are 'potentialities to M', where M is the manifestation of a given potentiality. The manifestation of a potentiality is "that property which the potentiality's possessor would possess if the potentiality were to be manifested" (2015, p. 102), so that it is sometimes called 'manifestation property'. Note that, as Vetter takes potentialities to be primitive, so the relation between a potentiality and its manifestation property is taken to be

[25] Note that a similar distinction can be done for essence and necessity, with essence being localized, being a property of a particular object, and necessity being non-localized, being a fact about how things are.

primitive by her as well, given that this relation is at the basis of the nature of potentiality itself: a potentiality is always a potentiality for a certain manifestation to be the case.

To clarify what she means by 'localized' and 'non-localized', Vetter appeals to the linguistic distinction between predicate and sentential operator (see again 2015, p. 2), suggesting that the kind of operator that ascribes potentialities is a *predicate operator* like '… (singular term) has the potentiality to… (predicate)', while the kind of operator that expresses possibility is a *sentential operator* like 'it is possible that… (sentence)'.

However, we must not confuse the distinction between localized and non-localized with the distinction between *de re* and *de dicto* modality, because while the latter applies to propositions, the former is a metaphysical distinction. Even if Vetter appeals to a linguistic distinction between operators, this is only for heuristic reasons, because she considers the difference between potentiality and possibility to be a metaphysical one, so that while the former is a property of singular objects, the latter is not.[26] Nevertheless, it seems quite clear that potentiality has the treatment of *de re* possibility statements as its most natural application, which is an issue that accompanies Vetter throughout her book and will need to be overcome when dealing with counterfactuals.

In any case, given the nature of potentiality as localized, and the fact that this distinction is primarily metaphysical, we can consider potentialities as *possibilities rooted in objects*, in the sense that "they are like possibilities, but they are properties of individual objects" (2015, p. 3). This fits well with her identification as a new actualist, because, given Vetter's belief that potentiality is rooted in individual objects and that metaphysical possibility can be accounted for entirely in terms of potentiality, this means that Vetter supports the idea that metaphysical possibility too is rooted in the individual objects of the actual world.

For Vetter, the thesis that all possibilities are rooted in individual objects is justified by our understanding of dispositions. Therefore, part of her work focuses on a new way to conceive dispositions supporting the idea that dispositions are the ground rooting possibility in the individual objects. Once this new conception of dispositions is established, Vetter's work focuses on the generalisation from dispositions to potentiality which she considers essential

[26] This is because Vetter does not take propositions to be objects in a substantial, Platonist sense, but only abstract objects, which in general she does not consider should be the bearers of potentialities (see 2015, pp. 270).

for two reasons. On one side, this generalisation allows potentialities too to be understood as the ground that roots possibility in the individual objects of the actual world. On the other, the generalisation is essential for a non-reductive realist metaphysics of dispositions, because, through potentiality, dispositions can gain a new metaphysical background without needing to be reduced to counterfactuals or possible worlds.

Given that her framework both requires dispositions as basis for the generalisation and offers a metaphysical background for dispositions, Vetter identifies her account primarily as a dispositionalist account. Here though, we will not focus on her work on dispositions directly, but only on those parts of it that are instrumental for understanding potentialism.

Her motivation for moving from dispositions to potentialities is that the traditional dispositionalist approach to possibility, in the shape of (P), can only cover a very limited portion of modal discourse. Consider (P) again:

> (P) "A possibility statement of the form 'It is possible that *x* is *M*' is true just in case *x* has a disposition to be *M*". (Vetter, 2015, p. 14)

As it is, (P) is formulated in terms that are not general enough because it can account only for a limited range of possibilities, those *de re* ones expressed by '*a* is *F*', while there are many *de dicto*, or more complex, or free-floating possibilities (e.g., the possibility that it is raining or sunny, the possibility that there are talking donkeys, etc.). In general, (P) simply has the wrong logical form to deal with these other possibilities because it covers only disposition ascriptions where an object has a *disposition to be or to do something*. Vetter's response then is to suggest a different logical form for disposition ascriptions, so that what is needed is "a general understanding of an object's having a disposition *for it to be the case that p*" (Vetter, 2015, p.17).

The move from dispositions to potentialities is meant to serve this purpose, because it is only through a generalization into potentiality that Vetter can make the notion of dispositionality wide enough so that it can account for dispositions (and so potentialities) "for it to be the case that *p*, for any value of *p*" (Vetter, 2015, p. 18).

The aim of her work is then to develop a notion of potentiality in a way that allows potentialities first, to be individuated by their manifestation only, rejecting both the conditional analysis (*x* is disposed to *M*, iff, if *x* were *S*, x would *M*) and (C) ('If *x* were *S*, then *x* would be *M*' is true just in case *x* has a disposition to *M* if *S*); second, to be potentialities 'for it to be case that

p' but maintaining that they are localized modalities; and third, to be so broad that they can accommodate for all kinds of (dynamic) possibilities.

Potentiality conceived in these terms differs from possibility overall because is relative to a particular object, so we may think of potentiality as possibility "relativized to a particular object" (Vetter, 2015, p. 18). However, this is not the direction Vetter wants to take, because she wants possibility to be defined in terms of potentiality and not vice versa: so, we should rather think of possibility as "potentiality *in abstraction from* its bearer" (*Ibid*.). Therefore, an improved version of (P) would be:

> (P') "It is possible that *p* just in case something has a potentiality for it to be the case that *p*". (*Ibid*.).

However, the final version of her account is actually:

> **POSSIBILITY** "It is possible that *p* $=_{df}$ Something has an iterated potentiality for it to be the case that *p*". (*Ibid*.)

This is the account of possibility we are interested in because it is the basis for Vetter's semantics, which includes the account of counterfactuals we want to discuss. We will not investigate fully how Vetter manages to reach such a definition of possibility, but we will see that the introduction of different kinds of potentialities, like iterated potentiality, which was originally intended for this purpose, is extremely relevant for her account of counterfactuals as well.

4. Vetter's Account of Counterfactuals

We have seen that, within the traditional account of dispositions, counterfactuals are the natural starting point because, if dispositional properties are taken to be individuated by a stimulus and manifestation, and these two are taken to relate to each other like the antecedent and the consequent of a counterfactual, then a semantics of counterfactuals could be developed in terms of dispositions in the shape of (C). However, we know that Vetter rejects this account of dispositions, supporting the idea that dispositions and potentialities are individuated by a manifestation only.

If there is no dyad of stimulus and manifestation involved in dispositions, then we get no straightforward semantics for counterfactual conditionals in terms of dispositions.

Counterfactuals require two elements to be considered for their account, an antecedent and a consequent, which are somehow related to each other. (C) can explain this relation in terms of stimulus and manifestation by linking the antecedent to the stimulus and the consequent to the manifestation, so that 'If *x* were to *S*, then *x* would *M*' is true iff *x* has a disposition to produce a certain manifestation *M* in response to a certain stimulus *S*. However, Vetter's potentialism cannot appeal to such a relation between stimulus and manifestation to reflect this feature of counterfactuals, because she only allows for one of these elements, the manifestation. Therefore, she needs another way to capture counterfactuals.

In what follows, we will just introduce Vetter's account of could-counterfactuals and some of her remarks about it, so that in our presentation there will be some concepts and solutions that have not been explained or justified yet. It will be matter of the next chapter to clarify the background of this account and to explore more in details Vetter's picture of potentiality in relation to counterfactuals, both in metaphysical and semantic terms.

Also, it is important to note that Vetter knows that what follows is only an initial suggestion for a potentiality-based semantics of counterfactuals, and so she is aware of the very much *tentative* nature of her proposal. She knows that her account comes with many unanswered questions, but she does not aim to answer them in her book. However, it is only through a deeper analysis of her tentative proposal that we can get a more substantial account, so this will be the aim of this work: to find out what sort of extra work and explanation are needed for Vetter's account to be acceptable.

Let us now see what this initial account is. Vetter models her proposal from the truth conditions she develops for 'can' statements, which are taken to work as ascriptions of potentialities. Such truth conditions are presented in (CAN):

(CAN) "'*x* can *F*' is true just in case *x* has a potentiality to *F*." (Vetter, 2015, p. 214)

Given this, Vetter starts not with would-counterfactuals, but with could- or might-counterfactuals, which she construes as ascribing iterated potentialities to individuals as well. Indeed, 'could' (and 'might' under a dynamic reading) can be considered a direct variant and analogue of 'can', so that could-counterfactuals can be more evidently taken to be ascribing potentialities, as 'can' statements do, and can be accounted for by an extension of (CAN).

The idea then is to take a counterfactual like "If I were taller, I could get that jar" to be ascribing *iterated potentialities*, where a potentiality – the potentiality to acquire the property

of being taller – in turn brings with it another contextually relevant potentiality – the iterated potentiality to get a particular jar. Following this interpretation, Vetter formulates (COULD), offering truth conditions for could/might-counterfactuals:

> **(COULD)** "'If *x* were *F*, then *x* could/might be *G*' is true iff *x* has an iterated potentiality to be *G*, and being *F* is an earlier stage in that iterated potentiality". (Vetter, 2015, p. 226).

After this definition, in the next few lines Vetter quickly suggests how to deal with other counterfactuals. Concerning past-counterfactuals, she claims that these "can be read as ascribing such iterated potentialities in the past" (*Ibid.*). Concerning cases where antecedent and consequent do not share their subject, these require "a simple generalization, where the iterated potentiality in question is a joint potentiality of their different subjects" (*Ibid.*). Concerning would-counterfactuals, these may be defined as "the dual of 'could' or 'might' counterfactuals: 'if it were that A, it would be that C' is simply shorthand for 'not: if it were that A, it might be that not C'" (2015, p. 227). All these cases are treated in an extremely superficial way in Vetter's book (quite literally by the quotes we have used above), so they will need to be discussed more in depth before we can accept that variations of (COULD) are effective in providing their truth conditions. This will be the focus of Chapter V.

However, before even trying to extend it to other counterfactuals, there are of course many points that need to be clarified concerning this proposal, starting by explaining what iterated potentiality is. To understand the metaphysical background of (COULD) we will need to discuss in some details the three main kinds of potentialities introduced by Vetter: joint potentialities, extrinsic potentialities, and, obviously, iterated potentialities.

Some clarifications are required in terms of semantics as well. For example, we will need to explain that Vetter's semantics, by working as an extension of (CAN), is compelled to take counterfactuals both as *context-sensitive* – "by requiring *x*'s potentiality to be of a certain minimal degree, or to be of a certain granularity, or perhaps to be of the right 'agentive' kind" (*Ibid.*) – and as expressions of *dynamic modality* only.

The restriction to dynamic modality explains why (COULD) treats 'could' and 'might' in the same way. On the assumption, that we will discuss better later, that Vetter's semantics is restricted to expressions of dynamic modality only, all the counterfactuals that she considers must be expressions of dynamic modality, so that (COULD) applies primarily to could-

counterfactuals, which are generally taken to be dynamic modals. However, Vetter accepts another assumption as well, which was introduced in Chapter I: that modal expressions are naturally ambiguous between different readings. This means that might-counterfactuals, under a certain reading, can be taken to be dynamic modals, despite 'might' being generally considered an expression of epistemic modality. Since Vetter accepts that might-counterfactuals can be dynamic modals as well, she considers that (COULD) can equally apply to those might-counterfactuals expressing dynamic modality.

Effectively, the difference between 'could' and 'might' as one being expression of dynamic modality and the other being expression of epistemic modality is ignored by Vetter here: she considers both 'could' and 'might' only in their dynamic reading, so that the two counterfactuals can be treated in the same way. However, Vetter points out that we must be careful in treating them as identical. Coming back to the example above, "If I were taller, I could get that jar", Vetter says:

> "A similar construal would apply to [the example] with 'could' replaced by 'might' (on a dynamic reading) – except, perhaps, that the conditions for contextual relevance at the second stage of the iterated potentiality are different: in the 'could' version, it is easy to read the sentence as requiring that the potentiality to [get the jar] would be an ability, while with the 'might' version that reading is less natural". (Ibid.).

All the clarifications concerning these points, from the different kinds of potentiality to context-sensitivity and the restriction to dynamic modality, will be done in the next chapter.

However, even if both the metaphysical and the semantic background surrounding this account have been extensively covered in it, Vetter's book does not offer much more insight on the treatment of counterfactuals themselves, and this leaves us with a sketch of an account that raises many questions that remain largely unanswered. In trying to fill these gaps and make of it a more substantial account of counterfactuals, we will see that there could be some serious threats to Vetter's potentialism. Indeed, I will argue that some attempts to develop the sketched account of counterfactuals could pose problems for the integrity of her picture as a whole. We will see that these problems are not only semantic but also, and more crucially, metaphysical. The conception of iterated potentiality suggested by this account, in the shape of the expression 'being an earlier stage', brings along certain perplexities that will need to be addressed. We will see in Chapter VI that the idea of 'iteration' in this sense is too vaguely defined to be applied to

counterfactuals in a straightforward way. Only a full understanding of this notion can help unpack her account and make it applicable to counterfactuals other than could-counterfactuals.

5. Conclusion of Chapter III

In this chapter we have started our discussion of Vetter's potentiality account of counterfactuals. In sections 1 and 2, we introduced dispositionalism in general and why a dispositional account of modality is preferrable to a possible world account. In sections 3, we presented Vetter's approach to dispositionalism and in section 4 we had a first taste of her account of counterfactuals.

Vetter's book focuses primarily on how dispositionalism, through her potentialism, can solve the problems arising from the limitations imposed by (P). In doing so, she introduces some essential concepts, like 'iterated potentiality', that she then employs in her account of counterfactuals. Our aim here is not to discuss how Vetter manages to turn (P) into POSSIBILITY, but it is important to remember that the notions used concerning counterfactuals, and that we will examine in the next chapter, were primarily introduced with this objective. This means that there is a lack of details about how these notions precisely work in relation to counterfactuals, and this work will try to investigate and raise questions on them, in an attempt to fill some of these gaps.

The next chapter will continue the discussion of potentiality, by focusing on its metaphysics and the semantics that is based on this metaphysics, both in relation to (COULD).

CHAPTER IV

COUNTERFACTUALS AND POTENTIALITIES

In the last chapter, we have introduced Vetter's version of dispositionalism and her proposal for an account of counterfactuals within this dispositionalism, which she calls '(COULD)':

> **(COULD)** "'If *x* were *F*, then *x* could/might be *G*' is true iff *x* has an iterated potentiality to be *G*, and being *F* is an earlier stage in that iterated potentiality". (Vetter, 2015, p. 226).

However, to understand such a proposal we need to explore a bit more in depth certain aspects of her metaphysical and semantic framework.

We can identify two crucial features of (COULD) which can have some important consequences for Vetter's treatment of counterfactuals, and which concern two different areas of discussion. On one side, (COULD) appeals to a specific kind of potentiality, iterated potentiality, as its metaphysical basis, so that Vetter's metaphysical picture of this kind of potentiality is essential to understand the metaphysical grounds of (COULD). The fact that this picture is quite unclear in some parts will have some important consequences for the account of counterfactuals. On the other side, (COULD) is modelled as an extension of (CAN), so that some of the characteristics and limitations that Vetter imposes on 'can' statements will apply to could-counterfactuals as well. This has some important consequences for the semantics presented by Vetter because, as we mentioned in the last chapter, this means that the counterfactuals she considers must be context-sensitive and must be expressions of dynamic modality.

Therefore, in this chapter, we will focus partly on the metaphysics of potentiality, in those aspects of it that are essential to understand Vetter's account of counterfactuals, and then on the semantics developed from this metaphysics, again mainly in those parts that are relevant for the discussion of counterfactuals. We will not challenge Vetter's general picture at this point, as our intent is to investigate an account of counterfactuals that accepts and is inserted into such a picture. Therefore, here we will not consider the possible criticisms or controversies raised by the potentiality account in general.[27] The intent of this chapter is purely expository, so to allow a better understanding of the metaphysical and semantic background that Vetter pairs with counterfactuals.

The structure of the chapter will be the following. In the first section, we will introduce the metaphysical background of Vetter's account by discussing three main kinds of potentiality. In the second and third section, we will discuss the semantics developed from this background. First, we will focus on the definitions of possibility and necessity given by Vetter and then we will discuss the two characteristics of (CAN) that can influence an account of counterfactuals: its context-sensitivity and its restriction to expressions of dynamic modality. The final section will conclude the chapter in preparation for the expansion of (COULD) to other counterfactuals in the next chapter.

1. Metaphysical Background

In understanding the metaphysical background of (COULD), we need to explain that there are three main kinds of potentialities identified by Vetter, which are introduced to expand the potentiality framework, so to cover more cases of possibility, but at the same time to maintain the idea that potentiality is localized: joint potentiality, extrinsic potentiality, and iterated potentiality. We will not discuss the details of this extension here, but we can explain why these kinds are needed for it to happen.

[27] For some reviews of her book, see Contessa (2016) and Leech (2017); for some recent criticism and discussion of her framework, see the series of essay in *Philosophical Inquiries* (2020), including Azzano & Borghini (2020), Bird (2020), Yates (2020), Giannini & Tugby (2020), McKitrick (2020), Wildman (2020), Wang (2020) and Vetter (2020).

One of the problems for Vetter is that potentialities, being localized modalities, are *prima facie* taken to be *intrinsic* properties of the objects of the actual world, but there are many possibilities that do not pertain only to the intrinsic nature of these objects, like the possibility for me to stand two metres apart from you, which depends on both your and my position in space. Therefore, Vetter needs to provide a mechanism for extrinsicality to be able to account for these possibilities, but she wants to do so without denying the nature of potentiality as localized. The introduction of joint and extrinsic potentialities is aimed at solving this conundrum, as we will see. Even if these potentialities are not directly involved in (COULD), they are essential for understanding how these truth conditions can be extended both to counterfactuals involving "extrinsic possibilities" and to counterfactuals with different subjects, because, as we know, these require "a simple generalization, where the iterated potentiality in question is a joint potentiality of their different subjects" (2015, p. 226).

Another problem is that Vetter, as we have explained, needs to justify how there can be potentialities 'for it to be the case that *p*', for any form of *p*. Only if we can make sense of these potentialities is her formulation of POSSIBILITY acceptable:

> **POSSIBILITY** "It is possible that *p* $=_{df}$ Something has an iterated potentiality for it to be the case that *p*". (*Ibid*.)

The introduction of iterated potentiality is meant to confirm the form of POSSIBILITY, since Vetter considers that iterated potentialities are 'potentialities *for*' rather than 'potentialities *to*'. Even if iterated potentiality is the crucial kind of potentiality for (COULD), it is important to remember that it was primarily introduced to allow the formulation of POSSIBILITY, so most of the discussion of this potentiality that can be found in Vetter's book has little to do with her counterfactual account. Nevertheless, only through a good understanding of what iterated potentiality is we can have a real grasp of the truth conditions presented in (COULD).

In what follows, we will discuss briefly what joint and extrinsic potentialities are, while we will analyse in slightly more detail what Vetter says about iterated potentiality.

1.1. Joint Potentiality

Joint potentialities are defined by Vetter as "potentialities that are possessed by a number of objects together or in combination" (2015, p. 107), which can be considered

'systems'[28] consisting of any number of objects that are joint possessors of potentialities. We know that individual objects have potentialities, but it seems that several objects taken together can have potentialities as well. For example: you and I have the joint potentiality to see each other; if I have the potentiality to play the piano and you have the potentiality to play the flute, then you and I together have the joint potentiality to play a duet (see 2015, p. 105).

It can happen that the potentialities that are possessed jointly by the objects in the system are different from the potentialities possessed by these objects separately. For example, if we consider a system to be constituted by a bullfighter and a bull during a bullfight, then the bull has, together with the bullfighter, a potentiality to get angry that is much more excited (i.e., at a much higher degree) than if this bull were on his own, because it is the purpose of the bullfighter to provoke him. The potentiality to get angry that the bull has alone then is different from the potentiality to get angry that he has when he is confronted by the bullfighter, because one is a potentiality to get angry to a lower degree, while the other is a potentiality to get angry to a higher degree. However, this second potentiality remains a potentiality concerning the bull only, because it is still the bull alone that has a potentiality get angry, and not the bullfighter as well, because, at the end of the day, it is the bull that is made angry by the bullfighter.

As we can see, the manifestation property of a joint potentiality can take different forms, being either a relation between its possessors (seeing each other), a plural property (playing a duet), or a property possessed by only one of the possessors (getting angry).

Plural properties are different from relations because they are not ordered, so that there is no difference between *a* and *b* having the plural property *P* and *b* and *a* having *P*, while relations are more precise in this sense, so that *Rab* is different from *Rba*; and because they can be borne by any number of objects[29], while relations can only hold between a fixed number of

[28] The idea of 'system' comes for example from Lewis (1997), where he uses the example of a system composed by a disc, having a yellow surface, that is disposed to produce a sensation of yellow in ordinary human observers and a perceptually ordinary human being that is disposed to receive a sensation of yellow, from objects having a yellow surface.

[29] The example of 'playing a duet' is not the best in this sense, but the equivalent 'playing music together' can include any number of players.

objects, as they always are *n*-places relations. However, these first two kinds of manifestation are related to each other and are easily explained:

> "When the manifestation of a potentiality consists in a number of object's jointly possessing a plural property, or standing in a certain relation, it is only natural that the objects which would together manifest the potentiality also possess the potentiality together" (2015, p. 109).

Therefore, both in the case of relations and plural properties, we can easily see how the objects having these kinds of manifestation properties possess the corresponding potentiality *together*.

The real problem is with the third kind of manifestation, when the manifestation of a joint potentiality consists in some, but not all, of the objects involved having the corresponding property: as we said, the bull and the bullfighter together differ from the bull alone in terms of their potentiality for *only one of them* to get angry, in the sense that they differ in terms of their potentiality for *the bull alone to get angry*.

However, a jointly possessed potentiality that is manifested in a property belonging to only some of the objects involved seems unacceptable, because, in general, potentialities are manifested by their bearers, since a potentiality's manifestation is a property that the bearer possesses. It seems then that a potentiality possessed jointly by the bull and the bullfighter should not manifest in a property that is possessed by the bull only, but in one possessed by him and the bullfighter as well. Therefore, we must assume that the bull and the bullfighter do possess a property *together* when manifesting their joint potentiality concerning the bull getting angry. To understand the manifestation of this potentiality then, we need to find out what this property is. The property in question must be either a plural property or a relation, given that must be possessed by or stand between all objects having the joint potentiality. However, assuming it is a plural property, the formulation of this plural property must be something like "the property of being such that the bull gets angry". Assuming it is a relation, instead, the formulation of this relation should be "the relation in which two things stand just in case one of them is identical to [the bull] and [gets angry]; again, the relation may be described as *being such that* [the bull gets angry]" (2015, p. 111).

Considering certain joint potentialities, it seems plausible to think that they somehow depend on the individual potentialities of their possessors. For example, a key and a door jointly have a potentiality to stand in a relation of opening because the key has shape S_1 and an *individual* potentiality to open locks with shape S_2 and the door has a lock of shape S_2 which has

an *individual* potentiality to be opened by keys of shape S_1. We could wonder then if a joint potentiality has any basis in the intrinsic nature of the objects possessing it.

However, Vetter claims that, even if there are cases in which the joint potentialities are grounded in the individual potentialities of their possessors, we must not conclude that we can straightforwardly reduce joint potentialities to individual potentialities. When some objects have a joint potentiality, it is not enough that these objects have *some* individual potentialities that "fit" together, because these objects could still have other potentialities which interfere with the manifestation of the joint potentiality. For example, consider the case in which the key still has a potentiality to open locks of shape S_2 but does not have a potentiality to open locks made of certain materials, like wax, and the door's lock, despite being of shape S_2, is made of wax: in this case, the joint potentiality will not manifest.

Therefore, for the objects to be guaranteed to have a specific joint potentiality, *all* their individual intrinsic potentialities must fit together. Even if it is very difficult to define what this "fitting relation" might be, it still seems plausible to think that "the joint potentialities of any number of objects are grounded in their individual potentialities taken together" (2015, p. 114). Vetter's conclusion then is that joint potentialities are intrinsic to their bearers when these are taken collectively, being a matter of how things stand with the *objects taken together*, like the door and the key, and not a matter of how things stand outside the objects involved.

Vetter makes one final point concerning joint potentialities: "Any objects might in principle jointly possess some potentialities, on the sole and trivial condition that they exist" (2015, p. 122). This is because we can think of potentialities that the bull and the bullfighter can have together even when they are separated, like having a potentiality to be far from each other, or a potentiality to be in different countries. For us to be able to think like this, it must be possible to ascribe joint potentialities to the bull and the bullfighter even if they are at the opposite ends of the universe, and so it must be possible to ascribe potentialities possessed jointly by objects that have no spatial contiguity or causal interaction between them. After all, we must concede that "all the objects that exist jointly have, at least, a potentiality to co-exist with one another" (*Ibid*.). I might not have any interaction whatsoever with the Emperor of Japan, but since we both exist, we do have some potentialities together, like the potentiality to never meet, or the potentiality to live the same number of years. Ultimately, even if, for some reason, we should not be sharing any other joint potentiality, still we would always have a joint potentiality to co-exist with one another, because we both exist in the same reality. Therefore,

the only necessary condition for me and the Emperor of Japan to have some joint potentialities together is that we both exist.

Here we have seen the three aspects of Vetter's discussion of joint potentialities that are essential for her theory of possibility. First, the fact that joint potentialities can have manifestation properties 'to be such that' is a step forward towards the formulation of POSSIBILITY. Second, the fact that joint potentialities should still be considered intrinsic is crucial to maintain the idea that potentiality is localized. Third, and finally, the fact that the only condition for two objects to have joint potentialities together is their existence increases exponentially the number of possibilities that can be explained by the potentiality account. These aspects of joint potentialities are also fundamental to understanding extrinsic potentiality.

1.2. Extrinsic Potentiality

Extrinsic potentialities are defined as those potentialities that "depend on aspects of the world outside the object that has them" (Vetter, 2015, p. 127). This seems to go against the initial idea of Vetter's program, which took all potentialities to be intrinsic of the object that possess them, being a kind of localized modality, and so a matter of how things stand with this object and not of how things stand outside it. However, Vetter suggests that there are extrinsic potentialities, and that they can be justified because they depend on joint potentialities. We talked about joint potentialities as a kind of relation in which objects can stand, so the idea is that, as relations in general can give rise to extrinsic properties, so joint potentialities give rise to extrinsic potentialities.

Coming back to the key-door example, Vetter's idea is that, when there is a joint potentiality for the key to open the door, this generates an extrinsic potentiality of the key to open the door. This potentiality is extrinsic because it does not depend on the intrinsic features of the key only, but also on the door's existence and some of its intrinsic features, because if the door ceased to the exist, or changed its lock, the key would lose this potentiality. Note that, in general, when a potentiality's manifestation is a relation to a *particular other object*, like the key's potentiality to open a particular door, Vetter takes this potentiality to be extrinsic. However, if this manifestation is simply about "standing" in a given relation to *some object or other*, such as the key's potentiality to open doors with a lock of a certain shape, then this potentiality is taken to be intrinsic, since it does not depend on anything other than the key.

In any case, Vetter believes that the key's extrinsic potentiality to open a particular door depends on key's standing in the relation to the door of "having a joint potentiality for one to open the other" (Vetter, 2015, p. 124). This scheme can solve the problem of explaining how extrinsic potentialities fit into a metaphysics that conceives potentialities as localized modalities, because it suggests that extrinsic potentialities are based on intrinsic but jointly possessed potentialities.

This systematic understanding of the relation between intrinsic and extrinsic potentialities, which suggests that intrinsic potentialities, which can be both individual and joint, are basic, and extrinsic potentiality are a by-product of them, makes extrinsic potentialities uncontroversial because it shows that they derive entirely from the intrinsic ones.[30]

The fact that Vetter takes extrinsic potentialities to be fully grounded in joint potentialities shows that accepting extrinsic potentialities is not in contradiction with the claim that potentialities are localized modalities:

> "Since full grounding is transitive, anything that grounds the possession of a joint potentiality, thereby grounds the possession of the relevant extrinsic potentialities; and anything that is grounded in an extrinsic potentiality's possession is grounded in the having of a joint potentiality" (2015, p. 133).

Therefore, if the possession of joint potentialities fully depends on the possession of the intrinsic potentialities of the objects involved, then the possession of extrinsic potentialities too depends entirely on the intrinsic potentialities of these objects, allowing us to maintain that potentialities, of any kind, are localized.

Admitting that there are extrinsic potentialities helps extending the reach of the notion of potentiality because they allow for an object *x* to have potentialities to be such that *p*, where

---

[30] There are some cases that could be considered problematic for this conclusion, like when there are *n* objects involved in a joint potentiality but there could have been *n*+1: you and I have a joint potentiality to play music together, but also you and I and John could have had the same joint potentiality. In this case, as the number of objects involved changes, so does the number of intrinsic potentialities that must be considered together. However, as we said, here we just wish to present Vetter's metaphysics rather than challenge it, so we will not discuss such cases.

p is entirely about objects other than *x*. For example, through extrinsic potentialities we can explain how the greengrocer of my village makes possible that I should buy an apple, because she has an extrinsic potentiality to be such that I buy an apple; or how a pianist makes possible that you should enjoy Mozart's Piano Concerto No. 24, because she has an extrinsic potentiality to be such that you listen to Mozart's Piano Concerto No. 24.

Even if both joint and extrinsic potentialities are not directly involved in the formulation of (COULD) we can see why they must be considered within a potentiality account of counterfactuals. Other than for explaining counterfactuals with multiple subjects, as Vetter suggests and as we will see in the next chapter, the fact that there can be extrinsic potentialities is clearly important for extending (COULD) to counterfactuals that involve situations that are external to the object involved, which we could call 'extrinsic possibilities', like "If John obtained a restraining order against Mary, he could be ten metres apart from her at all times". By accepting extrinsic potentialities, Vetter can account for counterfactuals of this sort by simply saying that the potentialities involved in them are extrinsic.

1.3. Iterated Potentiality

We have seen that the introduction of joint and extrinsic potentialities can extend greatly the scope of Vetter's theory of possibility, but without giving up the idea that potentialities are localized. Also, they can extend the range of (COULD) to counterfactuals including situations that are external to the object involved and to counterfactuals with multiple subjects.

Now it is time to discuss the kind of potentiality that is directly taken to be the metaphysical basis of (COULD), iterated potentiality. As we said, the introduction of this kind of potentiality is primarily needed for allowing the formulation of POSSIBILITY. Let us recall the first version of Vetter's account of possibility, which is (P'):

> (P') It is possible that *p* just in case something has a potentiality for it to be the case that *p*. (Vetter, 2015, p. 18)

The problem with (P') is that the potentiality involved in it is still a potentiality to do or to be something, and not a potentiality 'for it to be the case that …'. The discussion of joint and extrinsic potentialities only allowed for the manifestation property of a potentiality to be a 'being-such-that' property, not the potentiality itself. The definition of possibility given by (P') as it stands is not well-formed, because we still do not know how an object can be ascribed a

potentiality 'for it to be the case that *p*', for any logical form of *p*. Iterated potentiality is introduced by Vetter to make sense of these odd potentiality ascriptions, because she suggests that it is by iterating potentialities multiple times that we can reach a potentiality for it to be the case that *p*. We will discuss this briefly in what follows, but Vetter's conclusion is that an iterated potentiality has as its manifestation not the possession of some property (i.e., a manifestation property), but indeed *p*, as in "for it to be the case that *p*", for any form of *p*. I do not wish to investigate fully how exactly this happens, but, for our purposes, it is important to understand how iterated potentiality works.

Vetter defines iterated potentiality as "a potentiality for another potentiality, which might itself be for another potentiality, and so forth" (2015, p. 105):

> "Things have potentialities to possess properties. Potentialities themselves are properties. So, *prima facie*, things should have potentialities to have potentialities. And the latter potentialities might themselves be potentialities to have potentialities. So there is nothing to prevent things from having potentialities to have potentialities, or potentialities to have potentialities to have potentialities to have potentialities … and so forth. I will call any such potentiality an *iterated* potentiality." (Vetter, 2015, p. 135)

We can understand iterated potentiality better by considering some examples. I do not have the ability to play the piano, but I have the ability to learn, and so acquire, the ability to play the piano; therefore, I have an iterated ability to play the piano, which is an ability to have an ability to play the piano (i.e., a potentiality to have a potentiality to play the piano). Water does not have the potentiality to break but has the potentiality to freeze and turn into ice, which instead has the potentiality to break. Therefore, water has the potentiality to acquire the potentiality to break, so water has an iterated potentiality to break, which is a potentiality to have a potentiality to break. Note that this can also be expressed in terms of 'numbers' of iteration. If we take what could be considered a non-iterated potentiality as a once-iterated potentiality, then I have a once-iterated ability to learn to play the piano and a twice-iterated ability to play the piano, and the water has a once-iterated potentiality to turn into ice and a twice-iterated potentiality to break.

Crucial for a good understanding of iterated potentialities is to see what their manifestation is. This is also important for an argument we will consider in later chapters concerning the possible application of the manifestation relation as a basis for the 'earlier stage' relation introduced by (COULD). Vetter makes quite clear that, when we talk about iterated

potentialities, we must distinguish between their immediate manifestation and their ultimate manifestation.

The *immediate manifestation* of an iterated potentiality is *another potentiality*, so that the immediate manifestation of my iterated ability to play the piano (that is, my potentiality to have a potentiality to play the piano) is the ability (i.e., potentiality) to play the piano; and the immediate manifestation of the water's iterated potentiality to break (that is, the water's potentiality to have a potentiality to break) is the potentiality to break. On the contrary, the *ultimate manifestation* of an iterated potentiality is "the manifestation of the potentiality that it is a potentiality to have, or the potentiality that is a potentiality to have a potentiality to have … and so forth" (2015, p. 136). Thus, the ultimate manifestation of my iterated ability to play the piano is 'playing the piano' and the ultimate manifestation of the water's iterated potentiality to break is 'breaking'.

The nature of the ultimate manifestation of an iterated potentiality can help us determine whether this iterated potentiality is extrinsic or intrinsic: for example, if its ultimate manifestation is a relation in which the potentiality's bearer stands to a particular other object, then the iterated potentiality is *extrinsic* (e.g., my iterated ability to play my grampa's piano, which depends on its existence and on some of its properties, like 'being in tune'); if this manifestation is to stand in a relation to some object or other, instead, then the iterated potentiality is *intrinsic* (e.g., my iterated ability to play any piano, which depends only on my own skills).

In this way, we can also make sense of extrinsic and intrinsic iterated potentialities that are possessed jointly. If there are potentialities to have other potentialities, then there should also be potentialities to *jointly* have other potentialities with other objects, which are iterated joint potentialities. A door might have the potentiality to acquire the joint potentiality to be opened by a specific key, which is an extrinsic iterated joint potentiality. A door might also have a potentiality to acquire the joint potentiality, with keys of a certain shape, to stand in the opening relation, which is an intrinsic iterated joint potentiality.

Even if I do not want to explain in too many details how Vetter uses iterated potentiality to account for POSSIBILITY, still in doing so she needs to make an assumption that it is quite crucial for her account of counterfactuals.

This assumption is that there is *no limit* on the number of iterations between potentialities, in the sense that an object can possess potentialities that have been infinitely iterated, like my iterated potentiality to have a very distant descendant who lives on Mars (this iterated potentiality includes a very large number of iterations because it is a potentiality for me to have a child, who has a potentiality to have a child, who has a potentiality to have a child, who has a potentiality to have a child, and so on, for finally reaching a descendant of mine who has a potentiality to live on Mars). We will discuss why Vetter must accept such assumption of infinity in Chapter VI, when we will talk about the transitive nature of the iteration relation, but in any case, such an assumption is necessary for Vetter because it is only by unlimited repeated iterations that an object can have potentialities that concern objects entirely distinct from it (both in time and space) and therefore finally be a potentiality for *p*. To see how this works, we can consider two new examples, that expand from the previous ones (see 2015, p. 138).

A piano teacher possesses the ability to teach students how to play the piano. The manifestation of this ability consists in another individual acquiring an ability: a student, who could be me, learning to play the piano. The teacher then has "an ability to enter, with any student, into a joint potentiality to be such that the student acquires the ability to play the [piano]" (*Ibid*.), so that the teacher has an intrinsic, thrice-iterated ability, whose manifestation is that a student plays the piano. Therefore, the teacher has an iterated potentiality for it to be the case that I play the piano.

A freezer has the potentiality to turn water into ice, so that when this potentiality is manifested, a quantity of water freezes and acquires the potentiality to break. The freezer then has a potentiality to enter, with any appropriate quantity of water, into a joint potentiality to be such that the water acquires the potentiality to break, so that the freezer has an intrinsic, thrice-iterated potentiality, whose manifestation is that a quantity of water breaks. Therefore, the freezer has an iterated potentiality for it to be the case that the water in this ice-cube tray breaks.

These examples cover quite a limited number of iterations, but Vetter aims to apply similar reasonings to much more complex and extended sequences of potentialities, like those explaining the possibility for a human being to reach the end of the universe, or the possibility for Earth to be saved from climate change by my recycling a plastic bottle. In this way we can see that, if there is no limit on the number of iterations, iterated potentialities can truly be potentialities 'for *p*', so that their ultimate manifestation is not a property, but *p*. The reason for

this 'for *p*' formulation is that the ultimate manifestation of an iterated potentiality may concern an object completely separated from its possessor. The teacher does not have an iterated potentiality to learn to play the piano (because she can play already), but she has an iterated potentiality *for* a student to learn to play piano. The freezer does not have an iterated potentiality to turn into ice and break, but it has an iterated potentiality *for* some water to turn into ice and break. The manifestation of these iterated potentialities is not *playing the piano* or *breaking*, it is that *a student plays the piano,* or *some water breaks*.

The assumption of infinite iterations is essential for this result, but, as we will see, it is quite a dangerous assumption in terms of counterfactuals, because it results in (COULD) making true many counterfactuals that we would not consider true.

Nevertheless, iterated potentiality so conceived offers an effective extension of Vetter's account of potentiality, because it allows that an object *x* to have an iterated potentiality for it to be the case that *p*, where *p* is entirely about objects other than *x*. This extension is even more effective than the one offered by extrinsic potentialities because it has the right form for POSSIBILITY, and it can accommodate even better the localized nature of potentiality. On one side, the manifestation of an iterated potentiality does not even need to be in the form 'being such that *p*', because iterated potentialities are already potentialities '*for p*'. On the other side, iterated potentialities can be intrinsic and still have manifestations that do not concern their possessors at all, as we have seen from the examples.

Even if the introduction of iterated potentiality in Vetter's framework was intended primarily to formulate and justify POSSIBILITY, she uses this concept in her formulation of (COULD) as well. However, this proposal seems to imply something about iterated potentiality that is not necessarily implied by POSSIBILITY: that iterated potentiality comes in *stages*. We can make sense of how Vetter envisages iterated potentialities working in relation to counterfactuals by using some examples of counterfactuals showing what seem to be the different 'stages' in the iterated potentiality involved: e.g., "If I had learned how to play the piano, I could play the piano in this room", or "If this puddle of water were frozen, then it could break". However, as we will see, things are not as simple as they seem and the idea of 'stages' in an iterated potentiality is in truth quite difficult to define.

Therefore, counterfactuals raise a lot of questions concerning the nature of iterated potentiality and the concept of iteration, which could represent a threat to the picture of potentiality offered by Vetter. These questions will be considered in the next chapters.

2. Semantic Framework – Possibility and Necessity

Now that we have discussed the metaphysical background in which (COULD) is inserted, we should consider how Vetter builds her semantics, and where (COULD) stands in it. This discussion will be given more space than the one on Vetter's metaphysics because there are many aspects of her semantics that are relevant for a discussion of counterfactuals.

In this sense, it is essential to analyse Vetter's definitions of possibility and necessity, not only because they are at the basis of her semantics in general, but because, as we will see, on one side, a restriction on POSSIBILITY is necessary for formulating an account of would-counterfactuals, and, on the other, the original version of POSSIBILITY can be taken as basis for an account of *de dicto* counterfactuals[31].

Also, we need to discuss the semantics developed by Vetter for 'can' statements, because her semantics of counterfactuals is thought as an extension of it. Therefore, some of the features of her treatment of 'can' statements will be reflected in (COULD) with some crucial consequences, again concerning the treatment of *de dicto* counterfactuals.

2.1. Possibility

As we know, Vetter's idea is that when we speak of possibility within the potentiality framework, we should intend it as 'potentiality *in abstraction from its bearer*', so that "a possibility is a potentiality somewhere or other in the world" (Vetter, 2015, p. 197). Therefore, Vetter wants to define the possibility operator in these terms. We know that she chooses to express this idea by using an existential quantifier, in the shape of "*something* has a potentiality". Also, now we know that she formulates the account in terms of iterated potentialities because this kind of potentiality has the correct logical form, being a potentiality *for p*. Let remind us of POSSIBILITY once again:

[31] Note that Vetter might object that there are not *de dicto* (dynamic) counterfactuals, in the same way as she claims that there are not *de dicto* (dynamic) 'can' statements. We will discuss her position on *de dicto* possibility claims later and we will defend the existence of *de dicto* counterfactuals in the next chapters.

> **POSSIBILITY** "It is possible that $p =_{df}$ Something has an iterated potentiality for it to be the case that p." (Vetter, 2015, p. 197)

Note that, in terms of grammar, the two sides of POSSIBILITY are in the present tense, but they must be read as timeless. While metaphysical possibility is not subject to change over time, potentiality is, since we can gain or lose potentialities at different moments in time (see Vetter, 2015, pp. 186-187), so we need to read 'has a potentiality' as untensed. We could reformulate POSSIBILITY to make this evident:

> **POSSIBILITY*** "It is possible that $p =_{df}$ Something *has, had, or will have* a potentiality for it to be the case that p." (Vetter, 2015, p. 199)

However, for simplicity, we will just assume that POSSIBILITY must be read as POSSIBILITY*.

We can see that, by POSSIBILITY, the validation of the truth of the possibility claim on the left-hand side rests upon the presence of a witness that can verify the existential claim on the right-hand side. Therefore, the verification of a possibility claim is a matter of finding an object that can be the witness of the corresponding existential claim, which must be the bearer of the corresponding iterated potentiality.

In this way, a simple *de re* possibility claim like "It is possible that *a* is *F*" is simply a matter of the object involved (*a*) having the relevant potentiality (the potentiality to be *F*) because this is sufficient for it to make the case that something (*a*) has a potentiality for it to be the case that *a* is *F*. Take (1):

(1) It is possible that I fall asleep

This possibility claim is true if I have a potentiality to fall asleep, because, by having the potentiality to fall asleep, I have a potentiality to be such that I fall asleep, and so I can be the witness of the existential clam that something has an iterated potentiality for it to be the case that I fall asleep.

In the case of more complex possibilities, there can be multiple witnesses to the existential claim. Take the following *de re* possibility claim:

(2) It is possible that you and I play a tennis match

In this case, the potentiality ensuring the truth of (2) is a joint potentiality. However, POSSIBILITY is formulated with a singular quantifier, so that, even if the relevant potentiality

is our joint potentiality to play a tennis match, the witness for the possibility of our playing a tennis match cannot be the plurality of you and me. Vetter claims (2015, p. 198) that we could include joint potentiality in our account by adding "Something has, *or some things have*…" to POSSIBILITY, but she considers this is not really needed because, whenever some objects jointly have an iterated potentiality for *p*, then any one of these objects has an extrinsic iterated potentiality for *p*. Therefore, in general, any bearer of this extrinsic iterated potentiality can do the semantical job of being the witness to the corresponding existential claim, so that the witness for the truth of (2) is just either you or me: you, by having the extrinsic potentiality to play a tennis match with me, or me, by having the extrinsic potentiality to play a tennis match with you. Each of us is sufficient for verifying the existential claim that something (you or myself) has a potentiality for you and me to play a tennis match.

However, we need to remember that it is the appeal to joint potentialities that gives the metaphysical justification for the witnesses of the existential claim in cases like (2) because, metaphysically, we need to appeal to joint potentialities to ground the extrinsic iterated potentiality involved. The possibility that you and I play a tennis match might have you as its witness, since you have the extrinsic potentiality to play a tennis match with me; but this extrinsic potentiality is grounded in the joint potentiality, possessed by you and me, to play a tennis match together. The full metaphysical picture must appeal to joint potentialities, but the definition of possibility itself needs not to do so.

Given that extrinsic potentialities are based on joint potentialities, and that, as said above, joint potentialities can virtually hold between anything that co-exist, then extrinsic potentiality is extremely flexible. This flexibility means that there can be many irrelevant witnesses to any given possibility. However, this is not a threat to the definition given by POSSIBILITY, because the relevant witnesses for a possibility claim are always the most *basic* ones (see 2015, p. 199), metaphysically speaking. I might serve as a witness for the possibility that you are eating an apple, if I have an extrinsic potentiality to be such that you are eating an apple (e.g., because I have an apple tree in my garden), but the relevant potentiality is yours: it is your intrinsic potentiality to eat an apple. Therefore, in the first instance, *you* are taken to be the relevant witness.

However, this does not exclude that there can be many *relevant* witnesses for the same possibility claim. Take the following, still *de re*, possibility claim:

(3) It is possible that my great-grandson will be a pianist

In this case, I am the relevant witness in the first instance because, even if I do not have a great-grandson, and I will never have one, it is still possible that I should have a great-grandson who becomes a pianist. The relevant iterated potentiality is mine because I have a potentiality to have a child who has a potentiality to have a child who has a potentiality to have a son who has the potentiality to be a pianist. Therefore, it is me who makes the case that there is something (i.e., myself) that has an iterated potentiality for it to be the case that my great-grandson is a pianist. However, if I will do have a great-grandson, then he might be himself the witness to the existential claim, because, by POSSIBILITY*, it is possible that *p* if there is, was or will be an object with an iterated potentiality for *p*. Therefore, we clearly have two valid relevant witnesses in this case. This means that, even if Vetter considers that the relevant witnesses must be the most metaphysically basic ones, she admits that there can be multiple relevant witnesses to the same existential claim, so that my grandson's witnessing does not cancel mine.

Following POSSIBILITY* we can also explain the case of possibilities in the past. Consider the following *de re* past-possibility claim (4):

(4) It is possible that Erasmus should have been an atheist

In the case of my possible grandson, his existence is not necessary for (3) to be true, because I could be the witness of (3) in his place. In general, future objects are not essential to the truth of the existential claim, as long as we can find a relevant witness to it in the present (or the past). However, in the case of past possibilities, like (4), the relevant past objects are essential to the truth of the existential claim. The possibility for Erasmus to be an atheist cannot be explained by the iterated potentialities of any object existing in the present, because the relevant witness for the past-tensed existence claim (something *had* a potentiality for Erasmus to be an atheist) should be Erasmus himself, by having had the once-iterated potentiality to be an atheist (see, 2015, p. 202). There could be other witnesses (e.g., Erasums' parents), but, in any case, they can only be in the past.

Up until now, we dealt with examples of possibility claims that were *de re*, which are in general the easiest ones to account for in terms of potentialities. However, we can see that we can apply POSSIBILITY to *de dicto* possibility claims as well. Take the following three examples from Vetter (see 2015, p. 202):

(5) It is possible that there be a woman president of the US

(6) It is possible that there be a human space station on Mars

(7) It is possible that humans should have three legs instead of two

Despite being de *dicto* claims, the key to the truth of (5), (6) and (7) is still simply to find a witness for the relative existential claim. For (5) the relevant witnesses for example could be Michelle Obama, who has a potentiality to be a woman president of the US, and so a potentiality for it to be the case that there is a woman president of the US. For (6) one relevant witness might be an engineer who has a potentiality to build a space station on Mars.[32] Claim (7) is the most challenging, because it is not easy to identify the witness in this case. Vetter's suggestion (see 2015, p. 202) is that the witnesses should be our pre-human ancestors, who had an iterated potentiality for there to be human beings (i.e., a potentiality to have offspring with a potentiality to have offspring that is human) and who had *presumably* an iterated potentiality to have offspring that is human and has three legs. She adds "presumably" because, within her framework, what is possible and what is not depends on the way things actually are, so that we can only *presume* that our ancestors once had such an iterated potentiality. The same reasoning works for other *de dicto* possibilities of existence, like the possibility of there being talking donkeys or the possibility of there being unicorns.[33]

---

[32] Note that here the potentiality involved must be extrinsic, because it is based on a potentiality that must be jointly possessed with Mars.

[33] In general, the case of the possibility that it could exist something that does not actually exists is tricky for the potentiality account. Consider the following examples: the solar system has eight planets, but there might have been a ninth planet; even if they do not exist in nature, many think that there might have been unicorns. All these are unactualized possibilities of existence. While Michelle Obama does have a potentiality to be such that something is a woman President of the US, this does not work for the examples above, because nothing has a potentiality to be the ninth planet, or to be a unicorn. Vetter discusses the problem of unactualized possibilities of existence in 2015, pp. 267-273. In short, she suggests "to accept that there are unactualized *de dicto* possibilities of existence in the absence of a corresponding *de re* possibility" (*ivi*, p. 268), so we simply need to find out what the bearers of the relevant potentialities for these *de dicto* possibilities are. There are two options: either the bearers are the objects that have a potentiality to *produce* an object of the relevant kind (e.g., evolutionary ancestors of present-day animals, which had a potentiality to develop in a way to produce unicorns) or they are the objects having potentialities to *constitute* such an object (e.g., even if nothing in evolutionary history had potentialities for the evolutionary development of unicorns, still there are particles that have a joint potentiality to constitute a unicorn). There are some issues with this proposal, but we will not discuss them here. The potentiality account can also deal with possibilities of non-existence of contingent ordinary objects (see 2015, pp. 273-277): the possibility that I might not have existed is grounded by my parents who could have not had a child. Most of the more local possibilities of non-existence are then accounted by the fact that a contingent object will

The justifications of the various examples of possibilities presented above, other than showing the application of POSSIBILITY, can help explaining how (COULD) can deal with those sorts of possibilities as well. The idea quite simply is that, if we can find a witness to the existential claim corresponding to a certain possibility claim, then that claim does indeed express something possible. This can be quite useful in explaining counterfactuals involving past possibilities and possibilities *de dicto* because, as we will see, we can justify the possibility claims that occupy the place of antecedent and consequent of a past-counterfactual or a *de dicto* counterfactual by simply finding an object that has or had the relevant potentialities.

2.2. Necessity

After giving an account of possibility, we can get an account of necessity by means of interdefinability:

> **NECESSITY** "It is necessary that $p =_{df}$ It is not possible that not-*p*." (Vetter, 2015, p. 203)

So that:

> **NECESSITY** It is necessary that $p =_{df}$ Nothing has a potentiality for it to be the case that not-*p*.

As we did with POSSIBILITY, we can explain the application of NECESSITY by looking at two examples. Both concern the necessity of identity, but we can start with the simplest case:

(8) It is necessary that Hesperus is Hesperus

The potentiality account can easily show that (8) is true, because Vetter claims (see 2015, pp. 180-182) that nothing has a potentiality to be such that a contradiction holds, because nothing can have a potentiality for a contradictory manifestation, like a potentiality to be red-and-not-red. Therefore, nothing can have a potentiality for the contradictory manifestation that Hesperus is not Hesperus. There are cases in which an object can possess potentialities with

---

generally have developed in some way from other things or will be constituted by some other things. Those things would have had a potentiality for the object to exist if it did not, but they would also have had a potentiality for the object not to exist if it did. Again, we will not enter in the details of this discussion, but we can see that Vetter offers explanations both concerning possibilities of unactualised existence and possibilities of non-existence.

opposite manifestations, but these manifestations are never truly contradictory: I have an ability to walk and an ability not to walk, but I do not have an ability to walk-and-not-walk; I have a potentiality to start walking, whose manifestation involves, *first*, not walking and *then* walking, but I do not have a potentiality to be walking-and-not-walking-*at-the-same-time*. It is intuitively true then that nothing has a potentiality to be such that a contradiction holds, so we can imply that nothing has a potentiality to be such that Hesperus is not Hesperus.

Take now the more complex case concerning the necessity of identity:

(9) It is necessary that Hesperus is Phosphorus

Even in this case Vetter claims (see 2015, pp. 203-204) that we can clearly see that nothing has a potentiality to be such that Hesperus is not Phosphorus, because any idea that this could be the case is based on the difference between the names "Hesperus" and "Phosphorus" and their meaning, while potentiality is not a matter of language or meaning but of reality. It is an object that possesses the potentiality with a manifestation property such as 'being identical to Phosphorus' and this object is simply Hesperus. Therefore, there is no metaphysical difference between Hesperus having a potentiality and Phosphorus having it because they are the same object. So, there is no difference between something having a potentiality to be such that Hesperus is not Phosphorus and something having a potentiality to be such that Hesperus is not Hesperus. Therefore, as nothing has a potentiality to be such that Hesperus is not Hesperus, so nothing has a potentiality to be such that Hesperus is not Phosphorus.

The definitions of POSSIBILITY and NECESSITY, other than being fundamental for the development of Vetter's semantics, are quite important to understanding how we can move from the truth conditions for *de re* could-counterfactuals given by (COULD) to truth conditions that can apply both to would-counterfactuals (via a restriction of POSSIBILITY and a consequent restricted definition of IMPOSSIBILITY) and to *de dicto* counterfactuals (via an application of POSSIBILITY). However, this will be matter of the next chapter so for now we will stay on Vetter's semantics by considering her treatment of 'can' statements.

3. Semantic Framework – (CAN)

After defining POSSIBILITY and NECESSITY, we can see how Vetter uses these definitions to develop a modal semantics based on potentiality. This semantics is based on the idea that the "modal language is used to ascribe or deny potentialities to objects" (2015, p. 214)

and that the large part of these ascriptions of potentialities is performed by the predicate 'can'. As we know, if 'can' is used to ascribe potentialities, then the truth conditions for 'can' statements can be formulated as follows:

(CAN) "'*x* can *F*' is true just in case *x* has a potentiality to *F*." (Vetter, 2015, p. 214)

(CAN) is only the starting point for a modal semantics based on potentiality because, for this semantics to apply to other modal expressions, we need to generalize beyond the case of 'can'. As we said, (COULD) is part of this generalization, so that it is modelled on (CAN) and treated as an analogue of it. This means that the way in which Vetter characterises (CAN) and 'can' statements must extend to (COULD) and could-counterfactuals. We previously identified two features of (CAN) that are quite relevant for the treatment of counterfactuals: the fact that (CAN) must be intended as context sensitive, and the fact that (CAN) only applies to other expressions of dynamic modality, so that expressions of epistemic and deontic modality are excluded from Vetter's semantics. The repercussions of these features have different weight because, while context-sensitivity is a recognised feature of counterfactuals anyway, the restriction to dynamic modality, which for Vetter is synonymous of a rejection of *de dicto* dynamic 'can' statement, can be potentially damaging for her account of counterfactuals.

In what follows we will analyse these two features and their relevance for a counterfactual account. Concerning context-sensitivity, we will only discuss one out of the three conditions for context relevance identified by Vetter (degree, granularity, and agency). The application of these conditions to counterfactuals is quite straightforward and uncontroversial, so it is not of primary importance for our purposes. However, the presentation of the condition that Vetter calls 'degree' can be very useful to understanding a possible application of context-sensitivity presented in Chapter VI. The way in which the context selects the threshold on the degree of a potentiality has some parallels with a possible use of the context in limiting the transitivity of the iteration relation, as we will see in Chapter VI, so it is worth discussing it in some details.

Concerning the restriction to dynamic modality, in principle I do not reject Vetter's idea of separating dynamic modality from the others: it seems reasonable enough to distinguish between modal expressions which allegedly talk about "reality" directly (dynamic) and those which talk about our knowledge of it (epistemic) or the social constraints that we impose on it (deontic). I can see why she believes that they should be maintained separate, and I do not consider this restriction itself a threat to the account of counterfactuals I am trying to pursue.

The real issue for me is what should follow such a restriction in Vetter's opinion: the rejection of *de dicto* dynamic 'can' statement. As we will see, this rejection is not very desirable for an account of counterfactuals wanting to include *de dicto* counterfactuals, such as the one I aim to formulate.

3.1. Context-sensitivity

The context-sensitivity of 'can' is due to the large number of cases in which it is used: other than to ascribe abilities, it is also used to express that the conditions are right for their exercise, or just that there is a possibility for something to happen. One and the same 'can' statement can express all these nuances in different contexts. "I can drive" can be true if I learned to drive, but can be false if I do not possess a car or have a car in my immediate proximity; it can be true if I learned to drive and there is a car for me to drive, but it can be false if in this same situation I just finished my tenth shot of vodka and I am even struggling to walk. In this sense 'can' is context-sensitive: "it is used to express different things in different contexts." (Vetter, 2015, p. 215).

At the same time, clearly it is not the case that any potentiality can be ascribed by any 'can' statement in any context. For example, it would be very odd if "I can drive" ascribed to me the ability to drive just because I am a human being with functioning hands and feet and I am sitting in a car, and not because I actually learnt how to drive it. Therefore, (CAN) must be intended as contextually restricted, so that Vetter's general assumption is that (CAN) must be read as (CAN*):

> **(CAN*)** 'x can F' is true in a context C iff x has a potentiality to F which is relevant in C. (Vetter, 2015, p. 217)

As we said, Vetter identifies three conditions for a potentiality ascribed by 'can' to be relevant in a specific context: degree, granularity, and agency. However, the application of these conditions to counterfactuals is straightforward enough and it is not very interesting for our purposes, so we will not discuss them here (for Vetter's presentation of them see 2015, pp. 217-223).

Nevertheless, understanding how the condition of 'degree' works can be helpful for a discussion happening in Chapter VI, so we will try to present this condition in some detail. To explain what Vetter means by 'degree' we need to make a short digression that comes back to

dispositions. This will clarify even more the difference between dispositions and potentialities but will also show how the context works in selecting potentialities. This is the mechanism that will inspire, in Chapter VI, a possible appeal to the context in relation to counterfactuals, to solve the problem arising from the assumption of infinite iterations.

Dispositional terms like 'fragility' are by nature context-sensitive and vague: what counts as fragile in one context might be different in another (e.g., a chair in an antique shop or in a regular furniture shop), and there are borderline cases between what does and does not count as fragile (e.g., a plant-pot). If this is the case though, then we cannot rely on these terms to explain the underlying metaphysics, rather we must try to find which entities can be the context-independent background on which the instances of these vague terms and their context shifts operate.

Vetter's suggestion is that we can distinguish between the context-insensitive properties and the context-sensitive ascriptions by calling the latter "dispositions" and the former "potentialities" and that we should think of the context-sensitivity of dispositional terms as a matter of *degree*: "whether something count as fragile in a given context is a matter of *how* fragile it is" (2015, p. 21).

Therefore, to be suitable to underline ascriptions of dispositions, potentialities should come in degrees. In this way, having a disposition like fragility is a matter of having the corresponding potentiality, the potentiality to break, to a contextually sufficient degree. For Vetter, this means that "the potentialities which objects have outrun the dispositions that we are willing to ascribe to them" (2015, p. 22), so that, even if a chunk of gold will never count as fragile, still there is no deep difference between this and a champagne glass because both have the potentiality to break, only they have it at different degrees.

If potentialities come in degree, then they can be ordered in a *spectrum* going from minimal to maximal potentialities. The *minimal* degree of a potentiality is "the object's being just barely suited to show the manifestation at all" (2015, p. 95), so that a potentiality possessed to a minimal degree is one possessed by anything that can have, even if only in very remote situations, the manifestation property of that potentiality: the potentiality to break is possessed to a very low degree by a bridge, or a diamond, but it is still indeed possessed, because they can, after all, break. A potentiality is instead possessed to the *maximal* degree when the object lacks the opposite potentiality: something has the maximal potentiality to break if it lacks a potentiality not to break. Vetter's idea is that maximality is the equivalent of necessity:

> "A disposition is possessed to the maximal degree by an object just in case the object can do nothing other than manifest it; that is, just in case it has no potentiality not to manifest it. Necessity is the dual of possibility: it is necessary that *p* just in case it is not possible that not-*p*. A maximal potentiality to *F*, if it is to be analogous to a necessity, should be equivalent to the lack of a potentiality not to *F*." (2015, p. 86)

Between the minimal and the maximal potentiality to break there is the entire spectrum of things that can break, between which there are those things that would count as having the disposition 'fragility' in certain contexts: "[t]here may be no context in which we would call a chunk of gold 'fragile', but it too lies on the continuous spectrum of decreasing fragility" (2015, p. 101).

Vetter denies that there should be any natural cut-off points on this spectrum, and she rejects that we should impose some arbitrarily. When we go from more to less fragile objects,

> "there is no natural division insofar as their fragility is concerned […] [because] [n]o two objects in the series are so different to each other […] as to provide a non-arbitrary stopping-point between them" (2015, p. 83).

There are cut-offs produced by linguistic contexts, which are arbitrary, but they have a linguistic nature, not a metaphysical one. The cut-off needed here should be a division between what does and what does not have a potentiality to break in a purely metaphysical sense. And Vetter's position is that we cannot have natural cut-off points of this sort in a metaphysical spectrum like this one.

Therefore, the idea is that all the objects in the spectrum possess one property, the potentiality to break, at different degrees, and it is the context that sets different thresholds on this spectrum, defining the minimal degree to which the potentiality to break must be possessed for an object to count as 'fragile' in that context.

Vetter's idea is that '*x* can break' works mostly as '*x* is fragile' does, since both are ascribing to *x* a potentiality to break. Therefore, she thinks that '*x* can break' too is selective about the degree of potentiality that it ascribes. Both "My desk is fragile" and "My desk can break" are not true in all contexts because both require the desk to possess the potentiality to break at a certain minimal degree. Therefore, a 'can' statement sets the threshold on the degree of a potentiality in a certain context, just like a dispositional term does. However, '*x* can break' seems to have a lower threshold than '*x* is fragile', because we are inclined to accept "My desk can break" as true in many more contexts than "My desk is fragile".

This should make clear how the mechanism of context-sensitivity works in relation to the degree of a potentiality. We will suggest that a similar reasoning could be potentially applied in relation to infinite iterations, so it was quite important just to fix this point.

Now we can have a look at how the context-sensitivity of 'can' statements extends to counterfactuals. The fact that Vetter's extension of (CAN) must result in counterfactuals being context-sensitive is not a problem for her account, because this is a widely accepted feature of counterfactuals already. This can be shown by considering two examples from Lewis (1973b, p. 421):

(10) If Bizet and Verdi had been compatriots, Bizet would have been Italian

(11) If Bizet and Verdi had been compatriots, Verdi would have been French

These examples can be taken as evidence that the truth conditions of counterfactuals are context-sensitive, so that the possibilities that are considered when evaluating the antecedent are constrained by the context in which the counterfactual is asserted. Most views on counterfactuals incorporate some version of this idea, so Vetter has not added anything new in this sense. In truth, the fact that counterfactuals are context-sensitive can help Vetter sorting some issues out, as we will see.

Concerning the three conditions formulated by Vetter (degree, granularity, and agency), I simply accept what she says, that (COULD) too should be restricted "by requiring *x*'s potentiality to be of a certain minimal degree, or to be of a certain granularity, or perhaps to be of the right 'agentive' kind" (2015, p. 227). As we said, I do not consider these restrictions controversial, and they are not very interesting for our purposes. What is interesting is that counterfactuals seem to require an additional condition that must apply on the number of iterations between the potentialities involved in antecedent and consequent, so that the context should somehow restrict how many iterations there are between them. What is said above about the context-sensitivity of (CAN) in terms of degree can help in formulating such a condition, but this will be matter of discussion in Chapter VI. Therefore, let us move on to Vetter's discussion of dynamic modality.

3.2. Dynamic Modality

If the context-sensitivity of (CAN) does not represent a problem for (COULD) and it can actually be useful for a potentiality account of counterfactuals, Vetter's position on the kind of modality to which (CAN) must apply has much more serious consequences for (COULD).

As said before, Vetter's idea is that her semantics can extend to all expressions of dynamic modality, but only to them, because it cannot extend to expressions of epistemic and deontic modality. We defined dynamic modality in Chapters I and II, suggesting that it is the kind of modality concerning our reality, being "a matter simply of how things really are" (2015, p. 216). We also suggested that 'can' is generally used to express dynamic modality, because, even if it has deontic uses and can have an epistemic reading, compared to other modals, 'can' appears to be the primary expression of dynamic modality.

Nevertheless, Vetter suggests that, even if 'can' is the most common expression of dynamic modality, the semantics of (CAN) is easily transferred to other expressions that are constructed as analogues of 'can', like the modal verb 'be able to', the suffixes '-able, -ible' and the construction 'it is possible for… to…'. When these expressions express dynamic possibility, they can be given the same truth conditions of (CAN), with the right kind of construction in '…' (e.g., *x* is *F*-able):

> (*) '…' is true in a context *C* iff *x* has a potentiality to *F* that is relevant in *C* (2105, p. 224).

In this sense, 'could' in could-counterfactuals (and 'might', when it has a dynamic reading) is taken to be an analogue of 'can' as well, hence the formulation of (COULD) inspired by (CAN).

If the extension to other dynamic modals can happen quite straightforwardly, Vetter denies that a similar move can be done for expressions of epistemic and deontic modality. This may seem a limitation of her account, but Vetter's idea is that it is not really a disadvantage, because a unified semantics for all the different modalities is not so desirable. This rejection of a unified semantics is in contrast with a widespread position within possible world semantics supporting the idea of a unification of modality via restricted quantifications over worlds.[34]

[34] With some exceptions, like Chalmers (2002).

Vetter does admit that such a unification is attractive, but suggests that it is an oversimplification, because there is an undeniable deep difference between root modality and epistemic modality (see 2015, pp. 232-246).

As we know, root modality is taken to include dynamic modality and deontic modality and is separated from epistemic modality. Vetter's idea then is that this separation is so strong that trying to include both epistemics and roots within one semantics only is just reductive and simplistic. I will not discuss the details of Vetter's argument here, as what matters for me are just some consequences of it. Very briefly, she suggests that the difference between root and epistemic modality is simply too deep and unbridgeable to allow a valid formulation of a fully unified semantics of dynamic and epistemic modality, and therefore her separatist approach is preferable.

The only aspect of this argument that interests me is her claim that this deep difference partly resides in the fact that 'can' and its analogues, when expressing of dynamic modality, should be taken to be *predicate modifiers*, in the sense that they only take scope over the sentence's predicate (in "*x* can *F*" the scope of 'can' is only on the predicate *F*); while modal expressions of epistemic modality should function as *sentence modifiers*, in the sense that they take scope over the entire sentence (in "It may be that *p*" the scope of 'may' is on the sentence *p*).

This contrast has some very important consequences concerning *de re* and *de dicto* modality. Indeed, Vetter believes that one of the implications of the fact that dynamic modals like 'can' are taken to be predicate operators is that we cannot form *de dicto* dynamic statements: predicate modifiers cannot take scope over sentences because they always need a subject (it is always that 'something' can *F*), so that they cannot be taken to express *de dicto* modality, which is the sort of modality that attributes a modal status to sentences (e.g., It is possible that *p* is true).

This rejection of *de dicto* dynamic statements is supported also by an argument that Vetter takes from Brennan (1993). This argument is the so-called '*argument from scope*' (see 2015, p. 236-238). The idea is that, in sentences including epistemics, the scope of the subject is systematically ambiguous while, in sentences including roots, the subject always takes wide scope. Take the two examples:

(12) Every radio may get Chicago stations and no radio may get Chicago stations

(13) Every radio can get Chicago stations and no radio can get Chicago stations

(12) is ambiguous between two readings, a contradictory and a consistent one. The contradictory reading is that it is true that every radio may get Chicago stations, but it is also true that no radio may get Chicago stations, where the conjunction of these sentences is clearly a contradiction since they cannot be both true. The consistent reading instead is that it is possible that every radio gets Chicago stations, but it is also possible that no radio does, where the conjunction is consistent since "there may be opposing possibilities, as long as not both of them are actualized" (2015, p. 237). The contradictory reading is a *de re* reading, while the consistent one is a *de dicto* reading.

On the contrary, (13) is always contradictory, because the *de dicto* reading, which made (12) consistent, is not available for it. We must read (13) as a contradiction, which indicates that 'can', unlike 'may', cannot take scope over the whole sentence, and so allows only for *de re* readings. One way to explain this discrepancy is to take the epistemic 'may' to be a sentence modifier and the dynamic 'can' to be a predicate modifier.

This should go in support of Vetter's position that dynamic modals can only be predicate modifiers and therefore do not allow for *de dicto* readings, so that expressions of dynamic modality cannot be *de dicto*. This is quite a controversial position because we do seem to meet quite often uses of 'can' that are *de dicto* like "Someone can see us" (as "It is possible that someone sees us") or "Horses can run at 55 mph" (as "It is possible that horses run at 55mph"). Vetter justifies it by claiming that these examples are only apparently dynamic *de dicto* statements, because in truth their *de dicto* reading can happen only if the predicate 'can' in them is given an epistemic reading, and therefore is turned into a sentence modifier (see 2015, p. 229).

Vetter's conclusion then is that her semantics of (CAN) only applies to dynamic modals which must be expressions of *de re* possibility. These limitations imposed on (CAN) and 'can' statements both in terms of dynamic modality and *de re* modality are extended to (COULD) and to counterfactuals, so that (COULD) must be restricted to counterfactuals which express dynamic modality and must be *de re*.

The fact that Vetter's semantics can only apply to counterfactuals that are expressions of dynamic modality, naturally limits the scope of her account. However, this might not be a crucial issue for me, given that a unified semantics has never been the main objective of this work and my interest across it has always been on counterfactuals that are clearly expressing

dynamic possibilities. Even if it is appealing to think that counterfactuals should be all treated in the same way, I can see why Vetter supports a separatist semantics so my criticism will not focus on this point.

Indeed, the real potential issue for me comes from the fact that Vetter does not believe that we can have *de dicto* readings of 'can' when taken as expression of dynamic modality, because it must be a predicate modifier. Given this, the semantics of (CAN) can extend only to *de re* modality statements, where the modal expressions involved are constructed in analogy with 'can' as predicate modifiers. Therefore, my problem with Vetter's account is that (COULD) can only apply to counterfactuals in which 'could' and 'might' are taken to be analogues of 'can', and so predicate modifiers and expressions of *de re* modality. This is quite a drastic restriction to impose on an account of counterfactuals because I have assumed that there are *de dicto* counterfactuals, like "If it were the case that there are unicorns, it could be the case that there are flying horses as well", which in this way would just be excluded from Vetter's semantics.

She could object that, in truth, as there are no *de dicto* dynamic 'can' statements, so there are no *de dicto* counterfactuals that are *dynamic*, and if we think there are it is just because of an epistemic reading we can give to these counterfactuals.

In the next chapter, we will consider what could be her argument in favour of this position, but we will reject that this is the case, suggesting that indeed there are *de dicto* counterfactuals expressing dynamic modality. However, we will also show that they do not represent a threat to Vetter's view, because her account can be extended to this kind of counterfactuals, and we will suggest how this extension can happen.

4. Conclusion of Chapter IV

In this chapter we have explored in more details some aspects of Vetter's potentiality account of modality that can be relevant for an account of counterfactuals. In section 1, we discussed the three types of potentialities identified by Vetter, joint, extrinsic, and iterated potentialities. In section 2, we presented Vetter's definitions of possibility and necessity, which are at the basis of her semantics. Finally, in section 3, we analysed two features that Vetter's semantics ascribes to 'can' statements that she must ascribe to counterfactuals too, and that can influence (COULD): context-sensitivity and the restriction to *de re* dynamic modality.

We have seen that both the metaphysical and semantic pictures designed by Vetter have some repercussions on her treatment of counterfactuals, but it is crucial to remember, once again, that this background was not developed with a counterfactual account as its primary application, and that, because of this, there are many questions needing to be answered.

In the next two chapters, we will try to understand how these gaps can be filled, first by trying to extend (COULD) to different kinds of counterfactuals, and then by investigating the notion of 'being an earlier stage'.

CHAPTER V

A POTENTIALITY ACCOUNT OF COUNTERFACTUALS

The aim of this chapter is to further analyse Vetter's tentative account of counterfactuals, trying to see if and how it could be extended to counterfactuals other than *de re* one-subject could-counterfactuals (If *x* were *F*, then *x* could be *G*). In doing so, we will try to apply Vetter's semantics to past-counterfactuals (If *x* had been *F*, then *x* could have been *G*), counterfactuals with multiple subjects (If *x* were *F*, then *y* could be *G*), would-counterfactuals (If *x* were *F*, then *x* would be *G*), and *de dicto* counterfactuals (If it were that *p*, then it could be that *q*).

We will see that these extensions cannot all happen straightforwardly, and that Vetter's metaphysics needs a more extensive and precise formulation to accommodate all kinds of counterfactuals. In particular, the concept of "iteration" in the shape of the phrase 'being an earlier stage' is too vague to offer a sound metaphysical background for counterfactuals. The ambiguities in the definition of such a concept represent a common problem for all the extensions that I will suggest for (COULD), but my interest will be above all on the difficulties they create in the formulation of truth conditions for would-counterfactuals.

In what follows, I will accept Vetter's condition that her semantics applies only to counterfactuals expressing dynamic modality. The justification for such a position was given in the previous chapter. It can be controversial accepting that we cannot have a unified theory of counterfactuals that includes epistemic and deontic counterfactuals as well, and this could indeed be a reason to criticize Vetter's approach. However, it seems plausible enough to separate expressions that should allegedly concern our reality directly from expressions that

concern the "image" of reality coming from our knowledge or social conventions. Therefore, here I will not pursue this line of criticism against Vetter and will take the following counterfactuals as expressions of dynamic modality only.

Considering the second condition that Vetter imposes on her semantics, context sensitivity, we will discuss in the next chapter whether it can be of help in tackling certain issues. However, once again, we will see that the lack of a fully developed metaphysical background for the context to operate on will represent an obstacle for this solution.

After a quick first section on *de re* and *de dicto* possibility and impossibility, which will be instrumental for the discussion that will follow, the chapter will tackle, in the second and third sections, the simpler extensions of (COULD), past counterfactuals and counterfactuals with multiple subjects. The formulation of (WOULD), in the fourth section, will require more work since it should be the result of applying Lewis' interdefinability principle and translations between might- and would-counterfactuals. Since Vetter's proposal concerns primarily *de re* counterfactuals, our extensions will first apply to the *de re* versions of these counterfactuals, so that the discussion of *de dicto* counterfactuals will come last, in the fifth section. Once we have established if we can apply Vetter's semantics to *de dicto* could-counterfactuals, we will consider whether it could be applied to the *de dicto* variants of all the other counterfactuals in this chapter as well. The conclusion of the chapter will introduce the issues that will be discussed in the next one.

1. *De re*/*De dicto*

Before attempting to extend the scope of (COULD) to other counterfactuals, there is an important point that we need to make, to simplify our discussion. As we said in the last chapter, Vetter's account of counterfactuals addresses a kind of could-counterfactuals that is evidently *de re*, as we can see if we consider again (COULD):

> **(COULD)** "'If *x* were *F*, then *x* could/might be *G*' is true iff *x* has an iterated potentiality to be *G*, and being *F* is an earlier stage in that iterated potentiality". (Vetter, 2015: 226).

However, if we compare (COULD) with POSSIBILITY, we can see that this instead is clearly an account of *de dicto* possibility:

POSSIBILITY "It is possible that $p =_{df}$ Something has an iterated potentiality for it to be the case that p". (Vetter, 2015: 197).

Therefore, there is a sort of tensions between POSSIBILITY, which offers an account of *de dicto* possibility, and (COULD), which presents truth conditions for a *de re* type of counterfactual. The formulation of (COULD) in these terms is due to the fact that it is formulated starting from (CAN), and we said that Vetter considers that dynamic 'can' statements can only be *de re* modality statements. The definition of POSSIBILITY represents for Vetter a crucial and necessary attempt to extend her framework to the treatment of *de dicto* possibility, but, as we know, this extension is rejected for her semantics of (CAN) and, consequently, for her account of counterfactuals, so that her proposal in (COULD) only concerns could-counterfactuals expressing *de re* possibility. This could be a problem for a potentiality account of counterfactuals, as we will see in section 5, but for the moment we will accept this limitation. Therefore, our focus will be initially only on *de re* counterfactuals.

Because of this, for our discussion we do not need a *de dicto* definition of possibility, but it is sufficient a *de re* one. Therefore, for simplicity, we will adopt in this chapter a restricted version of "*de re*" POSSIBILITY[35] that can be roughly formulated as follow:

POSSIBILITY$_{\mathbf{dr}}$ It is possible for x to be $F =_{df} x$ has an iterated potentiality to be F.

This is obviously a limited version of Vetter's proposal, and one that does not fully respect her picture of possibility in relation to potentiality. Indeed, she admits that certain possibilities concerning an object might not be related to potentialities of the object itself, but to potentialities of other objects. For example, Vetter allows that the possibility for an individual to have been born in a different city may arise from potentialities possessed by other individuals, like the individual's parents, and not the individual themselves.

However, the formulation of POSSIBILITY$_{\mathrm{dr}}$ can be partly justified as follows. The POSSIBILITY definition of "It is possible for x to be F" is "Something has an iterated potentiality for it to be the case that x is F". As Vetter herself suggests (2015, p. 199), what are taken to be the relevant witnesses for a possibility claim like this are always the most *basic* ones,

[35] In the sense that we use '*de re*' here in a more restricted sense than usual, to cover only those cases of *de re* possibility (and impossibility) that do not depend on other objects but just on facts about the object of the *de re* modality statement.

metaphysically speaking. And in general, the most basic, and usually only needed witness of this potentiality is *x* itself, because *x* having a potentiality to be *F* makes the case that there is something having an iterated potentiality for it to be the case that *x* is *F*. Therefore, we could take POSSIBILITY$_{dr}$ as the definition of a certain kind of *de re* possibility that is restricted to those cases in which the relevant witness of the iterated potentiality must be the object of the *de re* statement itself.

For our discussion, the formulation of POSSIBILITY$_{dr}$ can greatly simplify things, especially concerning an extension of Vetter's account to would-counterfactuals. Since the would-counterfactuals considered will be *de re* counterfactuals, and they should be defined in terms of interdefinability with could-counterfactuals, as we will see, our proposal will need to go through a definition of *impossibility* that must be restricted in a similar way as POSSIBILITY$_{dr}$ to be applied effectively in this situation.

Considering POSSIBILITY, the correct definition of IMPOSSIBILITY should be the following:

> **IMPOSSIBILITY** It is impossible that *p* $=_{df}$ Nothing has an iterated potentiality for it to be the case that *p*.

However, a simplified and restricted version of this would be preferrable to develop an account of *de re* would-counterfactuals, so that again, we will adopt a restricted version of *de re* IMPOSSIBILITY:

> **IMPOSSIBILITY$_{dr}$** It is impossible for *x* to be *F* $=_{df}$ *x* does not have an iterated potentiality to be *F*.

Again, we can justify IMPOSSIBILITY$_{dr}$ by suggesting that this definition only convers those cases of *de re* impossibility where the witness of the iterated potentiality must be the object of the *de re* statement, so that for the statement to be impossible it is sufficient that this witness is missing. Regular *de re* impossibility statements requires that there is nothing that can pose as a witness to the iterated potentiality for it to be the case that *x* is *F*, and in this "nothing" is included *x* as well. Our proposal instead is restricted to those cases in which the simple fact that *x* is not the witness of the iterated potentiality makes the case for the impossibility of the *de re* statement.

This is a massive simplification of Vetter's account, but it is instrumental to allow for the formulation of an account of *de re* would-counterfactuals. Once this account is established, and

if we can extend it to would-counterfactuals in general, both *de re* and *de dicto*, then we can try to adopt again the original versions of POSSIBILITY and IMPOSSIBILITY as intended by Vetter.

However, before showing why and how much this simplification is beneficial for a formulation of the truth conditions of would-counterfactuals, we can consider two, more immediate, extensions of (COULD), concerning past-counterfactuals and counterfactuals with multiple subjects.

2. Past-counterfactuals

As we said in the last chapter, Vetter considers that past-tensed counterfactuals, like "If *x* had been *F*, *x* could/might have been *G*", should be read as ascribing iterated potentialities in the past.

When we introduced POSSIBILITY, we stipulated that it did not involve tenses or reference to time, and that it should be read timelessly. However, we have also seen that we can ascribe potentialities in the past, as indicated in the last chapter both by POSSIBILITY* and example (4) where Erasmus was taken as the witness of the past potentiality for Erasmus to be an atheist.

Therefore, the basic idea is that we can use counterfactuals to ascribe to objects potentialities concerning the past, and we can do so because those objects did possess those potentialities at some point in the past, either prior to the time of the counterfactual being asserted or prior to the time the counterfactual refers to.

We need to be careful here because we need to distinguish between the *ascription* of a potentiality, which is performed by the counterfactual at the time when this is asserted, and the *possession* of a potentiality, which happened at a time when the object indeed possessed such a potentiality to a certain degree. At the time of the ascription, the potentiality might be possessed at a zero degree, and therefore not be possessed at all, but this does not make the ascription false, because, for the ascription to be true, the relevant potentiality must have been possessed at least to a certain degree at the time to which the ascription is referring and not at the time when the ascription itself happens. I do not possess now a potentiality to have been in London on 1st January 2000, but this potentiality can still be ascribed to me now, because I did have a potentiality to be in London on 1st January 2000 before that date.

Vetter introduces a discussion of potentiality and time in her book (see Vetter, 2015, pp. 186-194) trying to explain why we are allowed to accept potentialities concerning the past. In there, she admits that it seems natural to think of potentialities as connected to time in an asymmetric way, so that the manifestations of potentialities seem to concern the future or the present, but not the past. However, she strongly supports the idea that there should not be a ban on past potentialities, because we do correctly ascribe them. We cannot exclude that I ever had a potentiality to be in London on 1st January 2000, because I did have it up until 1st January 2000. "Metaphysical possibility […] is a matter of something's possessing a relevant potentiality *at some time or other*" (2015, p. 187), so that, even if nothing has now a potentiality for me to be in London on 1st January 2000, it is still metaphysically possible even now that I could have been in London on 1st January 2000 because it is still true that I once possessed this potentiality. Therefore, Vetter suggests that we must reject a ban on potentialities with manifestations concerning the past.

I do not wish to discuss Vetter's treatment of past potentialities extensively, as my aim is to see how we can extend (COULD) to past-counterfactuals. Therefore, for the sake of our discussion, we will simply accept that past-counterfactuals can ascribe to objects potentialities concerning the past, on the basis that these objects possessed such potentialities in the past.

Now that this is established, we could try to adapt (COULD) for the case of past-counterfactuals. Recall (COULD), this time in a slightly extended version:

> **(COULD)** "If *x* were *F*, then *x* could be *G*" is true iff *x* has an iterated potentiality to be *G*, and a potentiality for *x* to be *F* is an earlier stage in an iterated potentiality for *x* to be *G*.

Now consider this first proposal for a past version of it:

> **(COULD-HAVE) (1)** "If *x* had been *F*, *x* could have been *G*" is true at time *t* iff *x had* an iterated potentiality to be *G* at some time t_1 prior to *t*, and, at t_1, a potentiality for *x* to be *F was* an earlier stage in an iterated potentiality for *x* to be *G*.

This first version of (COULD-HAVE) is intended for past counterfactuals in general, not referring to a specific time in the past, like:

(1) If I had been rich, I could have bought a bigger house

This counterfactual is true at *t*, which we assume is 'now', iff at some point t_1 in the past I had an iterated potentiality to buy a bigger house, and, at that time, a potentiality for me to be rich was an earlier stage in my potentiality to buy a bigger house.

Let us try to analyse these truth conditions. Considering the first conjunct, we assumed that we are entitled to ascribe past potentialities if the object involved did have such potentialities in the past: I did have a potentiality to buy a bigger house at some point in the past, because it is metaphysically possible that I should have bought a bigger house, so the first conjunct is satisfied. It does not matter if I still have a potentiality to buy a bigger house if, at some point in the future, I become rich: for the truth of the counterfactuals what matters is that I had this potentiality at the past time t_1. As said, whether I still possess or do not possess anymore the potentiality ascribed by the consequent when the counterfactual is asserted is not relevant for the truth of the counterfactual, because what counts is the time of the possession and not the time of the ascription.

Considering the second conjunct, in our proposal we are suggesting that the counterfactual is true if, at the time when I possessed an iterated potentiality to buy a bigger house, which is t_1, a potentiality for me to be rich was an earlier stage in this iterated potentiality. Once again, it does not matter if the latter potentiality would still be an earlier stage of the former now, because what matters is that this was the case at t_1.

As we can see, these truth conditions rely on the notion of a potentiality 'being an earlier stage' in another potentiality, so that for them to work we need to explain what this notion means. As we will discuss later, Vetter does not offer a complete picture of what it means for a potentiality to be an earlier stage of another, so that the details of this relation are missing, and we do not know what kind of consequences this connection has for the two potentialities involved. For example, we do not know whether there is some sort of dependence between them, or whether there is a strict unilateral temporal direction going from the earlier-stage potentiality to the later-stage potentiality, or whether both are the case, etc.

Even if the nature of this notion will be discussed extensively in the next chapter, we should investigate whether the use of 'earlier' implies some sort of temporal direction, because this could be relevant in our discussion of past-counterfactuals. Intuitively, 'earlier' seems to be used metaphorically here, to capture the fact that there must be some kind of "dependence" between the two potentialities involved as earlier and later stages, so that, roughly, the possession of the earlier-stage potentiality "causes" the possession of the later-stage one. The

temporal connection does not seem to be the primary meaning we should attribute to 'earlier'. Therefore, we will not pursue this line of reasoning. In any case, what counts for the formulation of (COULD-HAVE) is that it is at t_1 that the potentiality ascribed by the antecedent (call it potentiality *a*) must have been an earlier stage in the potentiality ascribed by the consequent (call it potentiality *b*), independently of whatever meaning we attribute to 'earlier stage'. This is because what matters for the truth conditions of the counterfactual is the time at which potentiality *b* was possessed, which must be some time in the past, and it is at this time that potentiality *a* must have been an earlier stage in potentiality *b*. Even if potentiality *a* is said to be an 'earlier stage' of potentiality *b*, this does not mean that it must have been possessed some time before *b* was possessed, because what counts for the truth of the counterfactual is that, at the time when *b* was possessed, *a* was an earlier stage in *b*, whatever this means.

In our version of (COULD-HAVE), neither the exact time of the antecedent nor the exact time of the consequent is specified by the counterfactual, so that, for the counterfactual to be true, it seems sufficient that the two potentialities involved, *a* and *b*, stand in a relation so that potentiality *a* was an earlier stage of potentiality *b* at the time when *b* was possessed, and they were both possessed at some time before the counterfactual assertion. We do not seem to need a specification of when the 'earlier-stage' potentiality *a* was possessed in relation to the possession of potentiality *b*, and so what sort of time lapse (if any) there should be between the two, because it is only important that this potentiality was an earlier stage of the other at t_1, whatever 'being an earlier stage' means.

However, things can get more complicated if the counterfactual itself includes a time specification, like:

(2) If I had been rich, I could have been in London on 1st January 2000

Indeed, it seems that my possession of a potentiality to be in London on 1st January 2000 should precede 1st January 2000. If on 1st January 2000 I was, say, in Australia, and there was no way for me to get to London by the end of day, then I could not have had a potentiality to 'be-in-London-on-1st-January-2000' on 1st January 2000. Therefore, it seems that the time specification requires that the possession of both potentialities should be prior to this specification.

(COULD-HAVE) (2) "If *x* had been *F*, *x* could have been *G* at t_1" is true at time *t* (after t_1) iff *x had* an iterated potentiality to be *G* at some time t_2 prior to t_1, and, at t_2, a potentiality for *x* to be *F was* an earlier stage in an iterated potentiality for *x* to be *G*.

Therefore, the counterfactual is true at the current time *t* iff at some time t_2 before 1st January 2000 I had a potentiality to be in London on 1st January 2000, and at that time t_2, a potentiality for me to be rich was an earlier stage in a potentiality for me to be in London on 1st January 2000. The necessity of t_2 as a time prior to 1st January 2000 is due to the fact, explained by Vetter, that the closer we get to the time in which the manifestation property of a potentiality will not manifest, the lower will be the degree in which the potentiality is possessed. Following her example (2015, pp. 189-190), if I am in Berlin this morning, I have a potentiality to be in Poland by 8pm today. A potentiality like this is extrinsic, because the degree of its possession depends on their possessor's positions in time: if I do not start moving towards Poland at a reasonable time, my potentiality to be in Poland by 8pm rapidly decreases in degree. If evening approaches, and I do not move towards Poland but remain in Berlin, my potentiality to be in Poland at 8pm gradually decreases in degree, reaching degree zero at 8pm, so that I no longer have the potentiality. Therefore, we can assume, for simplicity, that the time specification t_1 in a past-counterfactual actually corresponds to the moment in which the relevant potentiality turns to have zero degree or a degree that is too low to allow for the manifestation to happen (e.g., 23:59 on 1st January 2000) and so we must assume that this potentiality was possessed at some time prior to this moment. This is obviously a simplification, because it cannot be applied so straightforwardly to less narrow time specifications like 'last week' or 'last year'.[36]

Note that, whether the time specification is present in the counterfactual or not, all the past-tensed counterfactuals above are indeed past-counterfactuals, which are counterfactuals dealing with past possibilities. Even if the time is not specified, it is assumed that the use of the past tense in these counterfactuals is referring to a time in the past before the counterfactual

[36] In general, time specifications can be tricky for counterfactuals because some of them are sufficient to make them false, as in "If I had been rich at 23:59 on 1st January 2000, I could have been in London on 1st January 2000". However, this seems to be a general issue with dated counterfactuals rather than with Vetter's account. Vetter can easily explain why this counterfactual is false, by simply saying that a potentiality for me to be rich at 23:59 on 1st January 2000 was not an earlier stage of my potentiality to be in London on 1st January 2000, at whichever past time I possessed this potentiality.

assertion: it is in the past that I could have bought a bigger house. We point this out because sometimes the use of the past tense in counterfactuals does not necessarily set the counterfactuals at some time in the past: in a counterfactual like "If I had been a football fan, I could have enjoyed the match that I am watching now much more", the past tense is not used for referring to a past event, but rather for presenting a situation that is not the case. These counterfactuals represent a different case altogether, so we limit our application of (COULD-HAVE) to "real" past-counterfactuals.

Note also that there is no difference in the truth conditions of past-counterfactuals and those that we called 'historical' counterfactuals in previous chapters. The fact that the counterfactual is changing some well-known historical event is not relevant for any version of (COULD-HAVE). Take the following historical counterfactual:

(3) If Napoleon had been a better strategist, he could have won the battle of Waterloo

(3) is true now iff, at some time before the battle of Waterloo, Napoleon had an iterated potentiality to win the battle, and, at this time, a potentiality for him to be a better strategist was an earlier stage of his iterated potentiality to win the battle.

In this section, we have seen that we can extend (COULD) to past counterfactuals, but to understand how this extension work we need to establish whether the notion of 'being an earlier stage' really implies a fixed temporal direction. Above we assumed that it does not, because 'earlier stage' must not be interpreted in a temporal sense. Even if the notion of 'being an earlier stage' will be discussed better in the next chapter, we have already seen that intuitively it should be taken to correspond to some sort of "dependence relation" between potentialities, with the temporal direction only being the result of a linguistic metaphor. However, it is only through a clear definition of this notion that we can formulate some appropriate truth conditions for counterfactuals.

Naturally, what was introduced here is only an initial proposal of the treatment of past-counterfactuals and much more can be said about it. However, the extension I am mostly interested to is the one concerning would-counterfactuals, so I will leave this discussion for another occasion.

3. Counterfactuals with multiple subjects

Referring again to the last chapter, Vetter suggested that cases where antecedent and consequent do not share their subject require a "simple generalization", where the iterated potentiality in question is a joint potentiality of their different subjects. Let us have a look at how this generalization should work.

As usual, we start by (COULD):

> **(COULD)** "If *x* were *F*, then *x* could be *G*" is true iff *x* has an iterated potentiality to be *G*, and a potentiality for *x* to be *F* is an earlier stage in an iterated potentiality for *x* to be *G*.

Now, we can try to formulate the truth conditions for a counterfactual like "If *x* were *F*, then *y* could be *G*" by using extrinsic potentialities:

> **(COULD-*xy*) (1)** "If *x* were *F*, then *y* could be *G*" is true iff *x* has an iterated (*extrinsic*) potentiality for *y* to be *G*, and a potentiality for *x* to be *F* is an earlier stage in an iterated (*extrinsic*) potentiality for *y* to be *G* that is possessed by *x*.

This formulation seems simple enough and can be justified by what was said in the last chapter about extrinsic potentialities, where we accepted that an object could have extrinsic potentialities concerning entirely other objects. Take an example like:

(4) If Jane were an overachiever, then her teacher could be proud of her

This counterfactual is true iff Jane has an iterated extrinsic potentiality for her teacher to be proud of her, and a potentiality for Jane to be an overachiever is an earlier stage in Jane's iterated extrinsic potentiality for the teacher to be proud of her.

Let us analyse these truth conditions. Considering the first conjunct of (COULD-*xy*), we can see that Jane has an iterated potentiality for her teacher to be proud of her, because she has an iterated potentiality to create in someone else a feeling of pride and her teacher could be this someone. However, we also know that this iterated potentiality is extrinsic to Jane because it depends not only on her own potentialities and properties, but also on the potentialities and properties of the teacher, which is the person who should feel proud of her. As we know, the metaphysical justification of extrinsic potentialities relies on joint potentialities, so that Jane's extrinsic iterated potentiality for her teacher to be proud of her is grounded on Jane's and the

teacher's iterated joint potentiality to enter in the relation '*x* is proud of *y*', which, in turn, is grounded on the intrinsic potentialities of Jane and her teacher taken together.

Considering the second conjunct, we can see that again it contains the clause 'being an earlier stage'. Again, it is not clear what it means for a potentiality to be an earlier stage in another. However, the formulation in (COULD-*xy*) seems straightforward enough: the counterfactual is true if a potentiality for *x* to be *F* is an earlier stage in *x*'s iterated extrinsic potentiality for *y* to be *G*, so that (4) is true if a potentiality for Jane to be an overachiever is an earlier stage of Jane's iterated extrinsic potentiality for her teacher to be proud of her. It seems to be the case that the teacher could gain the property of being proud of Jane *following* Jane's success in her studies, and so that the two potentialities should be in some way connected on this basis. However, we are still relying on a very unclear picture of this connection.

If, when we talked about past-counterfactuals, we were interested in whether the phrase 'being an earlier stage' suggested a temporal direction between the two potentialities, in this case the only thing that is relevant for the truth conditions is that it implies some sort of consequential connection between them, so that the possession of the earlier-stage potentiality "causes" the possession of the relevant iterated potentiality. What we mean by 'causes' here though, still needs clarifying. We will investigate this idea first roughly when talking about would-counterfactuals below, and then more precisely in the next chapter, but at this point it seems intuitive enough to allow us to say that the second conjunct seems justified.

When dealing with multiple subjects, it is obvious that the potentiality ascribed by the consequent should be extrinsic, because it is an iterated potentiality concerning another object, which is the second subject of the counterfactual. However, we could question whether the potentiality ascribed by the antecedent, and so the earlier-stage potentiality, should be extrinsic or intrinsic to the first subject of the counterfactual. In our example (4) it seems intrinsic, since the potentiality to be an overachiever is apparently intrinsic to Jane. But does it need to be intrinsic?

No, because we can easily think of examples in which it is extrinsic:

(5) If Jane were 20m apart from my dog, my dog could still smell her

In this case, the counterfactual is true if Jane has an extrinsic iterated potentiality for my dog to smell her, and an *extrinsic* potentiality for Jane to be 20m apart from my dog is an earlier stage in Jane's extrinsic potentiality for my dog to smell her. Here, both potentialities involved

in the counterfactuals are extrinsic and need to be grounded in two different joint potentialities belonging to Jane and my dog: the joint potentiality to be in the relation of 'being 20m apart from each other' and the iterated joint potentiality to be in the relation of 'smelling'.

Our proposal in (COULD*xy*) is again just an initial suggestion and for simplicity it was formulated in terms of extrinsic potentialities, with a potentiality (either extrinsic or intrinsic) being an earlier stage in an extrinsic iterated potentiality. However, we could wonder whether our formulation should instead appeal directly to the metaphysical basis of this extrinsic potentiality: the joint potentiality possessed by the two subjects of the counterfactual. In this case, our proposal could look like this:

> **(COULD*xy*) (2)** "If *x* were *F*, then *y* could be *G*" is true iff *x* and *y* have an iterated joint potentiality to enter in a relation *R* such that if *Rxy* then *y* is *G*, and a potentiality for *x* to be *F* is an earlier stage in this iterated joint potentiality of *x* and *y*.

This is obviously a much more complicated set of truth conditions. Considering again (4), following this second proposal, the counterfactual is true iff Jane and the teacher have an iterated joint potentiality to enter in the relation of '*x* is proud of *y*' such that if the teacher is *x* and Jane is *y*, the teacher is proud of Jane; and a potentiality for Jane to be an overachiever is an earlier stage in an iterated joint potentiality of Jane and her teacher to enter in a relation of '*x* is proud of *y*' in those terms. These truth conditions become even more complex if also *x*'s potentiality to be *F* is extrinsic, as in (5), because we would need to introduce another joint potentiality in the picture.

This extra complication seems somehow unnecessary, given that extrinsic potentialities are grounded in joint potentialities anyway, but it gives the full metaphysical picture behind the truth conditions.

My conclusion is that it seems that the truth conditions in both versions of (COULD-*xy*) are formulated well enough, and the choice between them is only a matter of how extensive we want the metaphysical explanation to be. However, this does not imply that the metaphysics behind the notion of 'being an earlier stage' does not need to be clarified. But again, this clarification will be matter of discussion in the next chapter.

4. Would-counterfactuals

If the first two cases discussed above could be adapted from (COULD) without too much trouble, at least in term of 'structure' – since they still were could-counterfactuals – this is not the case for would-counterfactuals. As we know, Vetter clearly states that "'[w]ould' counterfactuals may be defined [...] as the dual of 'could' or 'might' counterfactuals" (Vetter, 2015, p. 227), suggesting that from (COULD) we should be able to construct an analysis of would-counterfactuals, which I will call '(WOULD)', based on this duality. Therefore, the formulation of (WOULD) cannot be constructed as a simple variation of (COULD) as (COULD-HAVE) and (COULD*xy*) but needs a more complex elaboration.

As we know, the existence of this duality is presented by Lewis, who supports the idea that might- and would- counterfactuals are interdefinable. Let remind us again Lewis' "translations" between the two:

I. "If it were that A, it would be that C" can be translated as "It is not the case that if it were that A, it might be that not C",

II. "If it were that A, it might be that C" can be translated as "It is not the case that if it were that A, it would be that not C".

Vetter's formulation of the account of counterfactuals in terms of could-counterfactuals is mainly due to its development from her semantics of 'can' as the primary modal predicate, and, since 'could' is the immediate counterfactual variant of 'can', it seems reasonable to treat it as an analogue of this predicate. Given her consideration of the predicate 'can' as a signal of ascription of a potentiality, the simplest and most natural way to approach counterfactuals in terms of potentiality is to initially look at those presenting a direct analogue of this predicate.

However, the traditional translations suggested by Lewis in *Counterfactuals* are between would- and might-counterfactuals, with no explicit mention of could-counterfactuals. As we saw in previous chapters, this led Vetter to opt for grouping 'could' and 'might' together, suggesting that could- and might-counterfactuals have the same truth conditions as given by (COULD).

We have seen that this move is legitimate because of the natural ambiguity of modal terms, so that 'might' can be given a dynamic reading. Someone could wonder whether it is justified to assume that Lewis' translations work for a dynamic reading of his examples, but it seems obvious that Lewis did conceive his might-counterfactuals as dynamic modals, given its

definitions of both might- and would-counterfactuals in terms of possible worlds, as seen in Chapters I and II, and the fact that he rejects Stalnaker's view of might-counterfactuals as only expressions of the speaker's epistemic status.

Therefore, Vetter is justified both in grouping could- and might-counterfactuals and in assuming that Lewis' translations can apply to her dynamic counterfactuals. In what follows, we will first show how these translations can help developing truth conditions for would-counterfactuals, and then we will test these truth conditions by applying them to some concrete examples.

At this point, we need to remind ourselves what we said about the difference between could- and would-counterfactuals in Chapter I: we said that the difference lies on the fact that, intuitively, the consequent of could-counterfactuals is *possible given the antecedent*, in the sense that is only one out of many possible outcomes given its antecedent, while the consequent of would-counterfactuals is *necessary given the antecedent*, in the sense that it must be the outcome of such an antecedent. This intuition needs to be represented by Vetter's semantics, and this representation will be crucial for the discussion below. Indeed, we will see that, to account for this difference, the notion of 'being an earlier stage' needs to be given a *modal* flavour so that we can interpret the possession of the earlier-stage potentiality as *making possible* the possession of the later-stage potentiality. First, though, let us see how the transition from (COULD) to (WOULD) takes shape.[37]

4.1. From (COULD) to (WOULD)

Recall once again our extended version of (COULD):

> **(COULD)** "If *x* were *F*, then *x* could be *G*" is true iff *x* has an iterated potentiality to be *G*, and a potentiality for *x* to be *F* is an earlier stage in an iterated potentiality for *x* to be *G*.

[37] The following two sections (4.1 and 4.2) have previously appeared in my article (Casini, 2022).

To proceed, let us take first the exemplifying could-counterfactual used by Vetter in (COULD) and turn it into the corresponding would-counterfactual, to which we then apply Lewis' translation:

(6) If *x* were *F*, then *x* would be *G*

(6*) It is not the case that if *x* were *F*, then *x* could be not-*G*

(6) is the counterfactual we are interested in and for which we want to find the truth conditions when formulating (WOULD). Hence, the construction of (WOULD) must start from (6*), because it is the negation of a could-counterfactual, whose truth conditions can be extrapolated from (COULD), but it is also equivalent to (6), so that its truth conditions will be those of (6).

There are some important remarks to make on (6*). Crucially, it involves negation, and in two instances: first, (6*) is a negation of a could-counterfactual; second, this could-counterfactual that is negated in (6*) has itself a negation in its consequent. Note that this second negation ("*x* could be not-*G*") takes narrow scope with regard to 'could', meaning "*it is possible for x not to be G*" (or "it is possible for *x* to be not-*G*"), and not "*it is not possible for x to be G*". This is quite an important distinction because the change of scope corresponds to a different treatment in potentiality terms. Also, here we can see the importance of our restricted versions of POSSIBILITY and IMPOSSIBILITY as $\text{POSSIBILITY}_{\text{dr}}$ and $\text{IMPOSSIBILITY}_{\text{dr}}$.

Considering $\text{IMPOSSIBILITY}_{\text{dr}}$, "*it is not possible for x to be G*" means that *x* does not have the potentiality to be *G*. The lack of a potentiality to be *G* by *x* is taken to correspond to the lack of the possibility for *x* to be *G*, and therefore to the impossibility for *x* to be *G*. Thus, the broad scope on the negation is taken to claim *x*'s lack of a potentiality to be *G*.

On the contrary, considering $\text{POSSIBILITY}_{\text{dr}}$, "*it is possible for x to be not-G*" means that *x* has the potentiality to be not-*G*. Rather than claiming a lack of a potentiality to be *G* by *x*, the narrow scope on the negation must be taken to mean the ascription to *x* of a potentiality to not be *G* (for a similar distinction concerning negation in Vetter, see her 2015, pp. 86, 145-146).

This is an important point in trying to determine the truth conditions of (6*), because it involves a narrow scope on the negation of 'could' and not a broad one. Therefore, our next step is to apply (COULD) to a counterfactual presenting such a negation:

(COULD~) "If *x* were *F*, *x* could be not-*G*" is true iff *x* has an iterated potentiality to be not-*G*, and a potentiality for *x* to be *F* is an earlier stage in an iterated potentiality for *x* to be not-*G*.

Now, to go from (COULD~) to the truth conditions of (6*) we should simply negate the former.

(TC6*) "It is not the case that if *x* were *F*, then *x* could be not-*G*" is true iff it is not the case that (*x* has an iterated potentiality to be not-*G*, and a potentiality for *x* to be *F* is an earlier stage in an iterated potentiality for *x* to be not-*G*).

Here though is where the first problem with Vetter's formulation arises. Like (COULD), (COULD~) is in the form of a *conjunction*: *x* has not-*G* AND *F* is an earlier stage of not-*G*. Therefore (TC6*) is the negation of a conjunction and as is well-known, the negation of a conjunction is in fact a *disjunction*:

(C/D) $\sim(A \wedge B) \equiv (\sim A \vee \sim B)$

As the initial truth conditions for could-counterfactuals – as per (COULD) – are conjunctive, then the negation of such truth conditions amounts to a disjunction of the negated conjuncts. Therefore, the truth conditions of the negation of a could-counterfactual are:

(~COULD) "It is not the case that if *x* were *F*, then *x* could be *G*" is true iff it is not the case that *x* has an iterated potentiality to be *G*, *or* it is not the case that a potentiality for *x* to be *F* is an earlier stage in an iterated potentiality for *x* to be *G*.

From which it looks like (WOULD) should be presented as a combination of (~COULD) and (COULD~):

(WOULD) "If *x* were *F*, then *x* would be *G*" ("It is not the case that if *x* were *F*, then *x* could be not-*G*") is true iff it is not the case that *x* has an iterated potentiality to be not-*G*, *or* it is not the case that a potentiality for *x* to be *F* is an earlier stage in an iterated potentiality for *x* to be not-*G*.

This formulation of (WOULD) seems promising, but to maintain the idea that a would-counterfactual is true if the consequent is *necessary given the antecedent*, we need to develop the concept of 'being an earlier stage' accordingly. A possible way to do this is through a *modal interpretation* of 'being an earlier stage', so that the fact that a potentiality to be *F* is not an earlier stage of an iterated potentiality to be not-*G* can be interpreted as the fact that *x* cannot

have a potentiality to be not-*G* if it has a potentiality to be *F*, and so *x* must have a potentiality to be *G* if it has a potentiality to be *F*. Here, I will tentatively suggest that this modal interpretation can be based on the existence or the non-existence of a chain of potentialities. To explain this idea, the best way is to look at the application of all the formulas above to some concrete examples.

4.2. Application: Colourful Cats

In formulating some examples to test the efficacy of our formulas, let us make three assumptions. First, that some ginger cats have a white belly – a ginger cat can be "white-bellied", while some other ginger cats do not have a white belly – a ginger cat can also be "non-white-bellied"[38]. Second, that all ginger cats are tabby. Third, that there is no white cat that is tabby.[39]

From these assumptions, if we take "Leo" to refer to a specific cat, we get an example for (COULD), "If Leo were ginger, he could be white-bellied", an example for (COULD~), "If Leo were ginger, he could be non-white-bellied", one for (~COULD), "It is not the case that if Leo were white, he could be tabby" and finally one for (WOULD) "If Leo were ginger, he would be tabby"[40].

Before applying our formulas to these examples, we need to make a final clarification: all features here, including 'being tabby' and 'being white-bellied', are taken to be exclusive and exhaustive with their opposite, like 'being non-tabby' and 'being non-white-bellied': all cats are either tabby or non-tabby, so that there is no cat that is both tabby and non-tabby and there is no cat that is neither tabby nor non-tabby.

---

[38] The use of 'white-bellied' is for maintaining the same "adjective" form of the other examples and formulas as "the potentiality to be white-bellied", "being white-bellied", etc.

[39] With this we mean an always completely and uniformly white cat.

[40] This example can flag another difficulty for Vetter, concerning the nature of iterated potentiality. We could doubt that the counterfactual "If Leo were ginger, he could be tabby" is really a matter of Leo's having two potentialities, one of being ginger and one of being tabby endowed by the first potentiality. We could instead think that there is just one potentiality: to be ginger in a certain way, namely in a tabby way. This is a worry raising from the ambiguous definition of iterated potentiality given by Vetter and the difficulties in understanding its stages, and it surely deserve attention, but for the sake of the argument we will assume for the time being that all the counterfactuals we are using are attributing two different potentialities to Leo.

Application of (COULD)

(7) "If Leo were ginger, he could be white-bellied" is true iff Leo has an iterated potentiality to be white-bellied (A), and a potentiality for Leo to be ginger is an earlier stage in a potentiality for Leo to be white-bellied (B).

As said, (COULD) is a conjunction, and a conjunction like (A $\wedge$ B) is true only if both A and B are true, so we can say that (7) is true only if both conjuncts in the application of (COULD) are the case.

Considering the first conjunct (A), the fact that Leo has an iterated potentiality to be white-bellied, simply means, by POSSIBILITY_{dr}, that it is possible for Leo to be white-bellied. Therefore, for (7) to be true, a necessary condition is that it is possible for Leo to be white-bellied, i.e., that the consequent is possible. However, considering the second conjunct (B), for the counterfactual to be true the other necessary condition is that a potentiality for Leo to be ginger is an earlier stage in the iterated potentiality for Leo to be white-bellied. My suggested interpretation of this condition is to take this iterated potentiality for Leo to be white-bellied as being part of a "chain" of other potentialities (and in this sense is "iterated"), the previous "chain rings" (or earlier stages) of which include a potentiality for Leo to be ginger. This idea of a chain of potentialities could be quite in line with our representation of could-counterfactuals as having the consequent *possible given the antecedent*, because if we take the existence of such a chain to which both potentialities belong as previous and later stages to mean that Leo's possession of a potentiality to be white-bellied is in some ways *made possible* (but not made necessary) by Leo's possession of a potentiality to be ginger, then we can clearly see that the consequent is made possible by the antecedent.

Application of (COULD~)

(8) "If Leo were ginger, he could be non-white-bellied" is true iff Leo has an iterated potentiality to be non-white-bellied (A), and a potentiality for Leo to be ginger is an earlier stage in a potentiality for Leo to be non-white-bellied (B).

This is another conjunction (A $\wedge$ B) so we can apply the same reasoning as before. Considering the first conjunct (A), once again the fact that Leo has an iterated potentiality to be

non-white-bellied means that, for (8) to be true, a necessary condition is that it is possible for Leo to be non-white-bellied (by POSSIBILITY$_{dr}$), i.e., that the consequent is possible. In the same way as before though, this is not sufficient, because for the truth of (8) the other necessary condition is that this iterated potentiality for Leo to be non-white-bellied is part of a chain having a potentiality for Leo to be ginger as an earlier stage, as per the second conjunct (B). Again, this could maintain the idea of the consequent of a could-counterfactual being *possible given the antecedent* because we could again interpret the existence of a chain to which the potentialities belong in previous and later stages as the fact that Leo's 'being non-white-bellied' is *made possible* (but again not necessary) by Leo's 'being ginger'.

Application of (~COULD)

(9) "It is not the case that if Leo were white, he could be tabby" is true iff it is not the case that Leo has an iterated potentiality to be tabby (~A), or it is not the case that a potentiality for Leo to be white is an earlier stage in a potentiality for Leo to be tabby (~B).

Here, in contrast with our previous examples, we are dealing with a disjunction. Therefore, we must keep in mind that for a disjunction like (~A $\vee$ ~B) to be true is sufficient that either one between the two disjunct is true, which suggests that it should be sufficient for the truth of (9) that only one of the disjuncts ~A and ~B is the case.

Considering the first disjunct (~A), the fact that it is not the case that Leo has an iterated potentiality to be tabby should mean, by IMPOSSIBILITY$_{dr}$, that it is not possible for Leo to be tabby, so that (9) should be true if it were *impossible* for Leo to be tabby, and this should be sufficient for the truth of (9). We could then interpret (~COULD) as saying that the negation of a could-counterfactual is true (and so a could-counterfactual is false) if the consequent is impossible. Note though that this impossibility is intended as unrestricted and independent of the antecedent, so we should clarify that, following the first disjunct, (9) is true if the consequent is *impossible simpliciter*, in contrast with *impossible given the antecedent*. With '*impossible simpliciter*' I intend a situation in which the consequent is "genuinely" impossible (and in this sense is "unrestrictedly" impossible) rather than a situation in which a normally possible consequent is not allowed by the antecedent, hence the expression '*impossible given the antecedent*'. Note that, obviously, if a consequent is *impossible simpliciter* it will also be

impossible given the antecedent: if it is impossible in every case, it will be impossible also when the antecedent is the case.

However, we know that it is not *impossible simpliciter* for Leo to be tabby, because it is metaphysically possible for Leo to be tabby, therefore we should consider the second disjunct (~B) for evaluating (9). In this case, it should be sufficient for the truth of (9) that a potentiality for Leo to be white is not an earlier stage in a potentiality for Leo to be tabby, which could be interpreted as the fact that there is not a chain of potentialities including both a potentiality for Leo to be tabby and a potentiality for Leo to be white as an earlier stage in this potentiality.

If we are entitled to interpret 'not being an earlier stage' as the fact that it does not exist a chain including both potentialities, and the non-existence of such a chain as the fact that Leo's possession of a potentiality to be tabby is *made impossible* by Leo's possession of a potentiality to be white, so that Leo's 'being white' makes it impossible for Leo to be tabby, then we get the desired and intuitive interpretation of the negation of a could-counterfactual, which is true not only when the consequent is *impossible simpliciter*, but also when the consequent is *impossible given the antecedent*.[41]

Therefore, it seems that these truth conditions for the negation of could-counterfactuals work if we interpret the negation of 'being an earlier stage' in a certain way. I will consider whether any questions might be raised about this interpretation in the next chapter so for the moment let us move to the application of (WOULD).

Application of (WOULD)

[41] Note that there is another case in which the negation of a could-counterfactual is true, as suggested by Vetter: "This [account] implies that the 'might/could' counterfactual is false when it is impossible for *x* to be *F*" (Vetter, 2015, p. 226, note 16). This can be explained in potentiality terms by saying that if it is impossible for *x* to have the earlier-stage potentiality to be *F*, then, if being *G* is part of a chain including being *F*, *x* cannot have an iterated potentiality to be *G* either, so that it is impossible for *x* to be *G* (first disjunct). Otherwise, if it is impossible for *x* to have a potentiality to be *F*, but *x* has a potentiality to be *G*, then the potentiality to be *G* must not be part of a chain including the potentiality to be *F* (second disjunct).

(10) "If Leo were ginger, he would be tabby" is true iff it is not the case that Leo has an iterated potentiality to be non-tabby (~A), or it is not the case that a potentiality for Leo to be ginger is an earlier stage in a potentiality for Leo to be non-tabby (~B).

Let us apply the same reasoning applied above, keeping in mind again that for the truth of a disjunction (~A ∨ ~B) is sufficient that only one between the two disjuncts is true, and so that it should be sufficient for the truth of (10) that at least one of the two disjuncts ~A and ~B is the case.

Considering the first disjunct (~A), the fact that it is not the case that Leo has an iterated potentiality to be non-tabby should mean that it is not possible for Leo to be non-tabby (by IMPOSSIBILITY$_{dr}$). Given our exclusive notion of 'tabby' and 'non-tabby', this means that (10) should be true if it were *necessary* for Leo to be tabby and this should be sufficient for the truth of (10). As before, we could then interpret (WOULD) as saying that a would-counterfactual is true if the consequent is necessary, but because this necessity is intended again as unrestricted and independent of the antecedent, it should be said that, following the first disjunct, (10) is true if the consequent is *necessary simpliciter*, rather than *necessary given the antecedent*. Once again, '*necessary simpliciter*' is used to indicate those situations in which the consequent is genuinely (unrestrictedly) necessary, while '*necessary given the antecedent*' indicates those situations in which a normally non-necessary consequent is enforced by the antecedent. And again, any consequent that is *necessary simpliciter* will be a consequent that is *necessary given the antecedent*, because if it is necessary in every case, it will be necessary also when the antecedent is the case.

However, once again we know that it is not *necessary simpliciter* for Leo to be tabby, because it is metaphysically possible for Leo to be non-tabby, so we need to consider the second disjunct (~B). In this case, it should be sufficient for the truth of (10) that Leo's potentiality to be ginger is not an earlier stage in Leo's iterated potentiality to be non-tabby, which could be interpreted as the fact that there is not a chain of potentialities including both a potentiality for Leo to be non-tabby and a potentiality for Leo to be ginger as an earlier stage in this potentiality.

As before, if we are entitled to interpret this 'not being an earlier stage', as the fact that there is not a chain including both potentialities, and the non-existence of this chain as the fact that Leo's possession of a potentiality to be non-tabby is *made impossible* by Leo's possession of a potentiality to be ginger, then we can say that Leo's 'being ginger' makes it impossible for

Leo to be non-tabby. Again, I will discuss whether any doubts might be raised about this interpretation but, if we accept it, given our exclusive notion of 'tabby' and 'non-tabby', then if it is impossible for Leo to be non-tabby, then it is necessary for Leo to be tabby. Therefore, we can interpret the second disjunct as saying that (10) is true if Leo's 'being ginger' makes it *necessary* for Leo to be tabby, so that we finally get the desired interpretation of a would-counterfactual, which is true not only when the consequent is *necessary simpliciter*, but also when the consequent is *necessary given the antecedent*.

We can see now that to be able to account for the difference between could- and would-counterfactuals in terms of the consequent being *possible* or *necessary given the antecedent*, Vetter's semantics must offer an interpretation of 'being an earlier stage' that is modal, so that a potentiality that is an earlier stage in another "makes possible" the possession of this second potentiality, while a potentiality that is not an earlier stage in another "makes impossible" the possession of it.

In my examples, I took that this modal interpretation depends on the existence or non-existence of a chain of potentialities connecting the two involved in the counterfactuals. If this interpretation is implemented, then we can see that by applying Lewis' translations and the rules of classical propositional logic we can reach an effective formulation of (WOULD). Naturally, this view needs to be justified and defended because the metaphysics behind it requires clarification. This will be subject of the next chapter, so for now we will have a look at *de dicto* counterfactuals.

5. *De dicto* counterfactuals

As we know, up until now all the counterfactuals discussed have been *de re* counterfactuals, so that our solutions have been limited to this kind of counterfactuals. However, I have claimed in previous chapters that I take it that there are *de dicto* counterfactuals as well, like "If it were the case that there are unicorns, then it could be the case that there are foals that are half horse and half unicorn".

We said in Chapter II that *de dicto* counterfactuals have the shape "If it were to be the case that the proposition *p* is true/false, then it could/would be the case that proposition *q* is true/false" like "If it were true that unicorns exist, it would also be true that there are horned horses". However, it is much more common to find *de dicto* counterfactuals leaving out any

reference to truth and falsity, like "If it were the case that unicorns exist, it would be the case that there are horned horses". What matters is that *de dicto* counterfactuals do not ascribe properties to specific objects, but rather investigate the consequences of the counterfactual truth or falsity of certain propositions. Nevertheless, it should be clear that I assumed that they could express dynamic modality.

In accepting that there are *de dicto* dynamic counterfactuals, we could face a possible issue with Vetter's semantics because, as said many times before, Vetter excludes that there are *de dicto* dynamic 'can' statements and extends this ban to all dynamic modality statements.

She suggests that, if we take 'can' to be a predicate modifier, as she does, then we cannot form dynamic *de dicto* statements with it because, given (CAN), every 'can' statement needs a subject. Vetter justifies her suggestion by claiming what seem to be some natural *de dicto* readings of certain dynamic statements are in truth not dynamic. She maintains that, in the case of dynamic statements like "Hydrangeas can grow on this soil" or "Someone can see us", which appears to have *de dicto* readings like "*It is possible that* there be hydrangeas growing on this soil" and "*It is possible that* someone sees us", these readings are actually not available if 'can' is taken to be dynamic, and if they seem available is only because of an epistemic reading of 'can'. Since 'can' read as an expression of dynamic modality only takes scope over predicates – being a predicate modifier – and not over sentences, it can only have a *de re* reading. To have a *de dicto* reading, 'can' should take scope over sentences and it can do so only if it has an epistemic reading and is taken to be a sentence modifier. Vetter's conclusion then is that we should reject dynamic *de dicto* 'can' modality statement.

Vetter's following step is to claim that we should deny that there are any *de dicto* dynamic modality statements in general (see 2015, p. 229, 232 and 237). However, while the rejection of a *de dicto* reading could be justified for dynamic 'can' statements, if we extend this to *all* other dynamic modality statements, and so we say that there cannot be a *de dicto* modality statement that is dynamic, this will make Vetter's semantics unapplicable to any *de dicto* counterfactuals. If *de dicto* counterfactuals cannot be taken as expressions of dynamic modality, and Vetter's semantics only applies to dynamic modality, then such counterfactuals will have to be excluded from this semantics.

Can we really accept that there is not any *de dicto* counterfactual which is an expression of dynamic modality? Vetter could argue in favour of this by using an argument similar to the one used for 'can' statement so that, as those dynamic 'can' statements that seem to express *de*

dicto modality are read this way only because we are giving an epistemic reading of 'can' in them, so those dynamic counterfactuals that seems to be *de dicto* are actually so just because of an epistemic reading that we give to them.

However, this strategy does not seem acceptable because we can easily think of some counterfactuals, like "If it were the case that there is a human with three hands, it would be the case that there is a human with fifteen fingers", which, despite clearly being *de dicto* counterfactuals, seem to express dynamic modality, in the sense that they consider an alternative way in which our reality could be, rather than talk about something depending on the knowledge of some speakers.

In truth, I consider that a rejection of dynamic *de dicto* counterfactuals is not acceptable for an account of counterfactuals like the one I want to pursue in this thesis, because I believe that there are indeed counterfactuals of this sort.

My conception of *de dicto* counterfactuals is that they can be taken to express dynamic modality. Therefore, if we want them to be included in Vetter's semantics, we need to accept that there can be dynamic modality statements which afford a *de dicto* reading. This option might not be as controversial as it seems. My idea is that, when she suggested that there are no *de dicto* dynamic modality statements, Vetter actually had in mind those statements that are most closely related to 'can' statement, without really considering the more extensive picture that includes counterfactuals. Indeed, we could say that what Vetter claims about *de dicto* readings is valid, but only for those expressions of dynamic modality that are construed in clear analogy with the predicate 'can', and therefore are construed as predicate modifiers (like the modal verb 'be able to', the suffixes '-able, -ible' and the construction 'it is possible for… to…').

In this way, even if 'can' and other similar expressions can be taken to be predicate modifiers, and so allegedly do not support *de dicto* dynamic readings, there might be other expressions that are not so clearly predicate modifiers that still can be expressions of dynamic modality and have a *de dicto* reading. Counterfactuals could be considered expressions of this sort. My interpretation of the counterfactual structure, as presented in Chapter I (If it were the case that *p*, it would/could be the case that *q*), seems much more similar to a sentence modifier than to a predicate modifier, because it does not apply to predicates, but to whole sentences, which are antecedent and consequent (i.e., *p* and *q*). Therefore, we could think that that this structure can be treated as an expression of *de re* modality, when antecedent and consequent are *de re* statements (If it were the case that *x* is *F*, it would/could be the case that *x* is *G*), but it

can also be treated as an expression of *de dicto* modality, when antecedent and consequent are *de dicto* statements (If it were the case that *p* is true, then it would/could be the case that *q* is true). However, this does not seem to be a reason to deny that in both cases counterfactuals can be expressions of dynamic modality: their first and primary reason to exist is for formulating hypothesis on how things could be different, and so they are clearly referring to the world itself and not to the image we have of it, whether they are *de re* or *de dicto*.

We can still derive the truth conditions for *de re* counterfactuals from (CAN), so we can still accept (COULD), (WOULD) and all the others, but we must agree that they can only work for *de re* counterfactuals. However, this does not imply that we cannot have *de dicto* counterfactuals that have dynamic readings. Simply, this means that the truth conditions of such *de dicto* counterfactuals cannot derive directly from (CAN), because these counterfactuals cannot be taken to be predicate modifiers but must be considered sentence modifiers.

If we take the counterfactual structure presented above (If it were the case that *p*, then it could/would be the case that *q*") as a sentence modifier, we can think of another case similar to a sentence modifier that is accepted by Vetter: possibility. As we know, POSSIBILITY is construed as a definition of *de dicto* possibility, as in "It is possible that *p*", for any form of *p*, so that it includes the case of *p* being a sentence. From POSSIBILITY we could formulate some truth conditions for *de dicto* possibility statements:

> **(POSSIBLE)** "It is possible that *p*" is true iff something has an iterated potentiality for it to be the case that *p*.

Since counterfactuals that are *de dicto* seem more similar to *de dicto* possibility statements than to 'can' statements in Vetter's picture, we could think of applying (POSSIBLE) to these counterfactuals. However, given that we are dealing with counterfactuals, we still need to explain the connection between antecedent and consequent, so it seems appropriate to maintain the same structure of (COULD), and the idea of 'being an earlier stage'. Take again:

> **(COULD)** "If *x* were *F*, then *x* could be *G*" is true iff *x* has an iterated potentiality to be *G*, and a potentiality for *x* to be *F* is an earlier stage in an iterated potentiality for *x* to be *G*.

In (COULD), the predicate 'could' is taken to be a direct variant of 'can' so that it is a predicate modifier which does not support a *de dicto* reading. Therefore, in considering the counterfactual structure for *de dicto* counterfactuals, we can apply (POSSIBLE) only if we adopt

an expression that is not a variant of 'can', nor a predicate modifier. What we need is an expression that can be taken to be a sentence modifier so that it can support *de dicto* statements. Expressions like 'it would be the case that' or 'it could be the case that', seem to be what we are looking for, because they are filled by sentences and so can support *de dicto* readings, but at the same time, the use of 'could' and 'would' in them helps to maintain the difference between could- and would-counterfactuals.

Therefore, we can formulate the following truth conditions for *de dicto* could-counterfactuals:

> **(De dicto)** "If it were the case that *p*, then it could be the case that *q*" is true iff something has an iterated potentiality for it to be the case that *q*, and an iterated potentiality, possessed by something, for it to be the case that *p* is an earlier stage in an iterated potentiality, possessed by something, for it to be the case that *q*.

In this case, we are truly appealing to iterated potentialities *for* something to be the case, and so we are applying the full extension of Vetter's account of possibility. Let us try to explain these truth conditions through a (slightly Austenian) example:

> (11) If it were the case that there are some bachelors in this room, then it could be the case that in this room there is a gentleman looking for an eligible young lady

This counterfactual is true iff something has an iterated potentiality for it to be the case that in this room there is a gentleman looking for an eligible young lady, and an iterated potentiality, possessed by something, for it to be the case that there are some bachelors in this room is an earlier stage in an iterated potentiality, possessed by something, for it to be the case that in this room there is a gentleman looking for an eligible young lady.

Let us consider the following scenario. The room has a potentiality to contain some bachelors (call it potentiality *a*). Between these potential bachelors, there could be at least one that has the potentiality to be in search of an eligible young woman (which could be considered an "embedded iterated potentiality", call it potentiality *b*). Therefore, we have a situation in which, on the condition that potentiality *a* is manifested – and so the room contains some bachelors – then potentiality *b* could be satisfied by at least one of these bachelors, who is looking for an eligible young lady. Which means that we can take potentiality *a* to be an earlier stage of potentiality *b*, because the fact that it is possible that the room contains some bachelors makes possible that the room contains a gentleman looking for an eligible young lady.

Within this picture let us consider each conjunct of the truth conditions. The first conjunct could actually be satisfied by any gentleman being in the room who is searching for an eligible young lady, who could be taken to be the witness of the relevant iterated potentiality *b* (remember that the first conjunct only establishes the possibility for the consequent to be the case). The second conjunct, however, adds a restriction concerning who this gentleman should be. It requires first that there is something which has an iterated potentiality for it to be the case that there are some bachelors in the room (potentiality *a*), and this something, as we know, could be the room itself which has a potentiality to have one or more unmarried men in it. But then this conjunct requires also that this potentiality *a* satisfied by any unmarried men in the room is an earlier stage in an iterated potentiality for the room to have a gentleman looking for an eligible young lady in it. And, indeed, this is the case in the scenario we described above, where at least one of the potential bachelors resulting from the manifestation of potentiality *a* could be the bearer of potentiality *b* and be a man in search for an eligible young lady.

Note that in this case, the counterfactual restricts the kind of gentleman that can be taken to satisfy potentiality *b*, because he must be one of the bachelors resulting from potentiality *a*. Therefore, case (11) is uttered to exclude those situations in which what satisfies one potentiality does not satisfy the other: e.g., the room having in it a bachelor (satisfying potentiality *a*) who is not in search for an eligible young lady (and so not satisfying potentiality *b*) or the room having in it a devious man who is in search for an eligible young lady (satisfying potentiality *b*) but is married (not satisfying potentiality *a*). For a Mrs. Bennet trying to marry off her daughters it is interesting that the room has in it some bachelors because they supposedly should satisfy potentiality *b*, while it is not interesting that the room has in it married men satisfying potentiality *b* because they do not satisfy potentiality *a*. Therefore, it seems that in this case, what satisfies both potentialities should be one and the same kind of object: a bachelor who is in the room and is in search of an eligible young lady.

However, in terms of the truth conditions, the 'something'-s satisfying potentiality *a* and potentiality *b* do not need to be the same object, even if most of times, including (11), this seems to be the case. Consider another example:

(12) If it were the case that there are unicorns, then it could be the case that there are foals that are half horse and half unicorn

In this case, the 'something' having an iterated potentiality for it to be the case that there are unicorns should be the horses' ancestors which had a potentiality to have descendants that

are unicorns, while the 'something' having an iterated potentiality for it be the case that there are foals that are half horse and half unicorn could be for example a horse which has a potentiality to breed with a unicorn. Therefore, it seems that the 'something' can be truly taken to be an expression of existential quantification, without any implication of this referring to the same object within the context of a counterfactual.

To try to extend these truth conditions to other *de dicto* counterfactuals, including past- and would-counterfactuals, we can see the following proposals. Note that counterfactuals with multiple subjects do not need a specific formulation of their *de dicto* truth conditions, since they can be included in (De dicto): *p* and *q* can be sentences with different subjects without affecting the formulation of (De dicto), because, as just said, the existential quantification does not require the 'something' to be the same object. Instead, both past counterfactuals and would-counterfactuals require a different formulation:

> **(Dd-COULD-HAVE)** "If it had been the case that *p*, it could have been the case that *q*" is true at *t* iff something *had* an iterated potentiality for it to be the case that *q* at t_1 prior to *t*, and an iterated potentiality for it to be the case that *p*, possessed by something at t_2 prior to (but extending to) t_1, *was* an earlier stage in an iterated potentiality for it to be the case that *q*, possessed by something at t_1.
>
> **(Dd-WOULD)** "If it were the case that *p*, then it would be the case that *q*" is true iff it is not the case that something has an iterated potentiality for it to be the case that not-*q*, *or* it is not the case that an iterated potentiality, possessed by something, for it to be the case that *p* is an earlier stage in an iterated potentiality, possessed by something, for it to be the case that not-*q*.

As you can see, with a formulation for *de dicto* would-counterfactual, in the first disjunct we can go back to adopting POSSIBILITY and IMPOSSIBILITY in their original *de dicto* form, because in saying that it is not the case that something has an iterated potentiality for it to be the case that not-*q*, we are saying that nothing has an iterated potentiality for it to be the case that not-*q*, and so we can apply IMPOSSIBILITY to say that therefore it is impossible that not-*q* and so that it is necessary that *q*.

All these formulations of truth conditions for *de dicto* counterfactuals are naturally just some initial proposals but they can help extending Vetter's account to several counterfactuals

that otherwise would be excluded. However, these proposals cannot do much for the main issue concerning the notion of 'being an earlier stage', which remains unclear.

6. Conclusion of Chapter V

In this chapter we have considered various attempts to extend Vetter's account to different kinds of counterfactuals. To do so, we first had to restrict Vetter's definitions of possibility and impossibility to some *de re* cases (section 1). We showed the utility of these restrictions in section 4, when talking about would-counterfactuals, which require a more complex elaboration than past-counterfactuals (section 2) and counterfactuals with multiple subjects (section 3). However, we saw that we can abandon these restrictions when we formulated a corresponding account for *de dicto* would-counterfactuals in section 5.

All our proposals seem quite plausible, but we can see that the concept of 'being an earlier stage' makes the extension of (COULD) to other counterfactuals difficult, because it is not explained nor developed enough to allow a straightforward application in other truth conditions, without requiring some extra work.

None of the work done in this chapter can be fully justified if we do not clarify this notion in a way that can metaphysically back up the various truth conditions put forward above. Therefore, the next chapter will aim to show what clarifications are required to the metaphysical picture offered by Vetter for making her account of counterfactuals acceptable. Whether these clarifications are indeed possible and how they should be done would require a kind of work that is probably too extensive to be included in this thesis but pointing out the existence of this ambiguity is already crucial for Vetter, because it could have serious consequences for her potentiality account.

CHAPTER VI

COUNTERFACTUALS AND ITERATION

Across this thesis, we have seen that counterfactuals seem to require an account of modal metaphysics because they are sensitive to metaphysical possibilities that can exceed what is taken to be possible within the actual world and therefore there are difficulties in accounting for them straightforwardly in terms of actuality. We have seen then that an account of metaphysical modality can explain them either by exceeding actuality as well, by appealing to non-actual possible worlds (modal realism) or by trying to keep these possibilities anchored to actuality, by reducing what is metaphysically possible to what is actual (actualism). Counterfactuals have been included in these explanations, both in terms of possible worlds and in terms of dispositions. Barbara Vetter wants to offer an account of metaphysical possibility based on actuality, by appealing to certain properties of the objects of the actual world, which are potentialities. Within her account of possibility, Vetter offers a tentative account of counterfactuals based on a specific kind of potentiality, iterated potentiality.

The aim of this chapter is to investigate the relation between iterated potentiality and counterfactuals and to discuss some of the problems arising from this relation when we analyse Vetter's tentative account of could-counterfactuals.

In the last chapter, we formulated some truth conditions for counterfactuals other than *de re* one-subject could-counterfactuals, like past-counterfactuals, counterfactuals with multiple subjects, would-counterfactuals, and *de dicto* counterfactuals. We saw that for the truth conditions of all these counterfactuals to work it is essential to explain the notion of 'being an earlier stage', because it is common to all these truth conditions, but it is left unexplained by

Vetter. In this chapter we will discuss this notion, and that of iteration, trying to find out what clarifications are needed for it to offer a solid account of counterfactuals.

Since Vetter does not offer an extensive explanation of the 'being an earlier stage' notion, what this chapter aims to do is to consider some possible suggestions that could arise from her intuitive picture. I will not attempt here to provide a conclusive answer, but I will evaluate some pros and cons of some proposals.

Note that here the main issue is not to try to understand what iterated potentiality is. Even if it would be interesting to explore this matter, the exact nature of iterated potentiality is not of primary importance for our purposes. What matters to us is that iterated potentiality is taken to be the metaphysical background of counterfactuals, so that the fact that *x* has a potentiality to be *F* which is somehow connected to an iterated potentiality for *x* to be *G* makes it true that "If *x* were *F*, *x* could be *G*".

Therefore, the chapter will not so much investigate what kind of entities iterated potentialities are, rather it will focus on the 'being an earlier stage' relation and so on how a potentiality can be connected to an iterated potentiality. To do so, we will introduce some options that are compatible with Vetter's presentation of iterated potentiality in relation to counterfactuals.

This 'earlier stage' relation seems to have some intuitive general features that can be easily recognised: it is asymmetric, transitive and, possibly, non-reflexive. We will see that the transitivity of this relation can create a big issue for counterfactuals, because the possibility of unlimited iterations seems to enforce the truth of counterfactuals that we would normally consider false. In analysing this problem, we will discuss whether an appeal to the context could help, but we will have to conclude that, for this to work, the metaphysics behind the context must be established more precisely first.

Another general feature of the 'earlier stage' relation has been identified in the last chapter, when dealing with would-counterfactuals: that this relation needs to have some modal consequences. To account for such an interpretation, first we will analyse our initial proposal concerning the existence of a chain of potentialities and then we will enquire whether we could interpret this relation directly as a purely modal relation. We will conclude that neither suggestion can work and that the best solution is actually to look for ways to characterize the 'earlier stage' relation in such a way that it can be given a modal interpretation.

Therefore, in what follows, we will discuss two main proposals on how to characterize this relation in a metaphysical way: on one side we can take the potentialities to be connected by a 'chain of repetitions', while on the other we can consider them linked via a 'chain of yielding'. Both views have their advantages and disadvantages. Here we will try not to favour one over the other, but we will discuss the merits and costs of both, since the decision on which view to adopt depends on many factors.

One of these factors is how much we value the extension of Vetter's account compared to its effectiveness and clarity. Vetter must assume certain things about iterated potentiality for it to play the role of metaphysical background of counterfactuals, and we will see that, to be able to account for *all* counterfactuals, Vetter seems to require an incredibly general notion of 'being an earlier stage', which could be in the shape of a chain of repetitions. In this sense, we will investigate whether it could be expressed more clearly in terms of the manifestation relation. However, even if this notion with this interpretation could help extend the account to all counterfactuals, it may be too general to be effectively explained.

In truth, we know that to account for *all* counterfactuals is a difficult issue for any account, because trivial counterfactuals or counterfactuals where the consequent does not seem to depend on the antecedent represent a problem. I clarified in Chapter I that in this work I am mostly interested in finding a metaphysical background for "regular" counterfactuals, rather than willing to discuss the exceptions, so that my research focuses those counterfactuals clearly involving some kind of *dependence* between antecedent and consequent.

Therefore, another factor is whether we accept or not this assumption, that counterfactuals should involve this sort of dependence of the consequent on the antecedent. If we do accept it, then we need Vetter to accept that there must be some kind of *dependence* between the potentialities that are connected in a counterfactual as well, because only through this dependence she can account for the dependence between antecedent and consequent, and this dependence could be explained in terms of a chain of yielding.

However, this approach could limit the extent of Vetter's account only to counterfactuals clearly showing a relation of dependence between antecedent and consequent, leaving out all others. Also, the main problem with this solution remains that the nature of this 'dependence' relation is not clear. To attempt to clarify it, we will explore different kinds of dependence relation and see if they could work in relation to a potentiality account of counterfactuals: could it be a relation of causal dependence? Or could it be characterized in terms of grounding? Once

again, within this 'dependence discussion', I will not commit to any of these answers, but I will discuss the problems that each could have.

My conclusion will be that the picture of iterated potentiality presented by Vetter on this matter is essentially too vague to work and it seems to create some confusion between the nature of iterated potentiality as potentiality to have a potentiality in a sequence of repetitions on one side, and the fact that 'fewer times'-iterated potentialities seem to "yield" 'more times'-iterated potentialities on the other. By consequence, much more work is needed before an account of counterfactuals could be built in terms of iterated potentiality.

The structure of the chapter will be the following. The first step in our investigation will be establishing, in the first section, why Vetter decided to use iterated potentiality for explaining counterfactuals and how this explanation could be characterised in two different ways: either as a chain of repetitions, or as a chain of yielding. The second section will describe more precisely the idea of a process of iteration, as the process connecting 'fewer times'-iterated potentialities to 'more times'-iterated potentialities, and how this could be described within both views above. In the third section, we will define the general features of the 'earlier stage' relation that are common to both views and will see that it is an asymmetric, transitive, and non-reflexive relation, with its transitivity being a potential issue for Vetter's account. In the fourth section, we will discuss the fact that, for providing an effective background of counterfactuals, this relation must have a modal nature, or at least modal consequences. The fifth sections will investigate three possible candidates for explaining this relation, one concerning the chain of repetitions and the other two relating to the chain of yielding: manifestation, causation, and grounding. The conclusion of the chapter will sum up the problems that Vetter must face concerning counterfactuals.

1. Iterated potentiality and Counterfactuals

The relation between iterated potentiality and counterfactuals was not investigated fully in Vetter's book, since her main objective in introducing iterated potentiality was to extend the scope of potentiality to *de dicto* possibility, rather than to apply it to the specific case of

counterfactuals. However, we can easily retrace what led Vetter to use this notion to explain counterfactuals.[42]

If we accept the very rough simplification presented in Chapter I, counterfactuals can be taken to be expressions of a relation existing between antecedent and consequent, where a certain consequent is taken to be a necessary or a possible outcome of a certain antecedent, so that such a consequent is said to be necessary or possible *given such an antecedent*.

What Vetter needs to do is to offer a metaphysical picture that can map such linguistic expressions and therefore give account for such a relation. As we know, traditional forms of dispositionalism could rely on a stimulus-manifestation relation to provide this metaphysical basis, but Vetter strongly rejects this kind of dispositionalism, so she needs to find another way to fit counterfactuals in her framework.

The introduction of iterated potentiality via the expression 'being an earlier stage' is intended to precisely play this role: it accounts for the fact that antecedent and consequent of a counterfactual are related to each other through the fact that the two potentialities ascribed by the counterfactual are related to each other as well, with one being the earlier stage in the other, which is iterated. The relation between an iterated potentiality and its earlier stage then is taken as the metaphysical basis of the relation between a consequent and its antecedent in a counterfactual.

This seems a plausible solution, where the reasons for the truth of a counterfactual in the semantics are backed up by the underlying metaphysics. However, what does it mean exactly for a potentiality to be 'the earlier stage' in another, iterated, potentiality?

Vetter does not explicitly explain it, and her language on this is always metaphorical, talking about a potentiality "bringing with it" another (2015, p. 226), or "being part of a chain of potentialities" (2015, p. 137). Even if there is indeed a sense in Vetter's presentation and language in which the possession of iterated potentialities is the result of the possession of its earlier stages, the nature of this relation is not clear.

In general, the idea that iterated potentiality comes in stages was introduced very quickly by Vetter and only in relation to her tentative account of counterfactuals: this idea is

[42] Parts of this section have appeared in my article (Casini, 2022)

presented only in the two pages (2015, pp. 226-227) where she introduces her account of counterfactuals, which is, by her own admission, only a sketch (the phrase "earlier stage" itself has only three entries in her entire book and is never accompanied by a definition).

What seems quite clear is that the idea of a potentiality 'being an earlier stage' in an *iterated* potentiality must be connected to the fact that this former potentiality is 'iterated'. The being-an-earlier-stage phenomenon then seems to be connected to the idea of 'iteration'. However, this does not solve our problem, because Vetter does not provide a full explanation of the concept of 'iteration' either, always maintaining it more metaphorical than substantial.

Therefore, in what follows we will investigate some possible explanations of the notions of 'being an earlier stage' and of 'iteration' that seem in line with her metaphorical language.

1.1. Two kinds of chain

An iterated potentiality is defined by Vetter as "a potentiality to have a potentiality" (2015, pp. 135-139). Therefore, we may take two views on what an 'earlier stage' could be, which we can characterize in terms of 'chains' connecting different potentialities.

On one side, an earlier stage could be a potentiality that is part of the *chain of repetitions* of the 'have a potentiality' function. Take a potentiality to learn to play the piano as being an earlier stage of a potentiality to play the piano: the once-iterated potentiality to learn to play the piano could be taken to relate to a twice-iterated potentiality to play the piano, which is a potentiality to have a potentiality to play the piano, by being the first repetition of the clause 'have a potentiality to', so that 'have a potentiality to learn to play the piano' corresponds to the first repetition of 'have a potentiality to' in 'have a potentiality to (1) have a potentiality to (2) play the piano'. In this case, an earlier stage is simply a potentiality that occurs earlier in a series of repetition of the 'have a potentiality' function (and hence, if we consider that potentiality is forward-looking in time, whose manifestations will be earlier than the manifestations of the later stages; more on that in section 5.1.)

On the other side, an earlier stage could be instead a separate potentiality that is part of a *chain of yielding*, in the sense that is a potentiality that yields the iterated potentiality, so that the once-iterated potentiality to learn to play the piano yields the twice-iterated potentiality to play the piano without being part of the repetitions in 'to have potentiality to have a potentiality

to play the piano'. In this case, an earlier stage is a potentiality that is distinct from the chain of repetitions of the 'have a potentiality' function but gives rise to them.

In the first case, 'iterations' are intended literally as '*repetitions*' of the 'have a potentiality' function, while in the second, 'iterations' are intended as 'fewer times'-iterated potentialities that *yield* other 'more times'-iterated potentialities.

We can find arguments in support or against both views. The 'chain of repetition' view has the advantage of being relatively simple and of linking the notion of 'earlier stage' to an extremely general relation that can cover a wide range of cases, without really needing any reference to the idea of dependence. However, this relation, being so general and simple, could also be considered too generic, in the sense that, for example, it does not provide enough content for the antecedent. If an earlier stage corresponds simply to a repetition of 'have a potentiality to' then how can we attribute any "real" content to it? Also, it remains to be seen whether leaving out the idea that there is a dependence relation between the potentialities is truly an advantage, since I assumed that such a relation should hold at least for standard cases of counterfactuals.

The 'chain of yielding' view instead, has the advantage of trying to be more precisely defined and avoiding the problems of being too generic, and it does try to account for the idea of dependence that I support as a basis for standard counterfactuals. However, this relation seems unable to cover the same wide range of cases of the other one, and it seems to exclude a line of explanation that is available for the other, a parallel with the manifestation relation that we will discuss further on in section 5. Indeed, by assuming that the relation between "stages" in an iterated potentiality should be one of dependence, this view blocks what could be considered the natural move of understanding the relation between stages in an iterated potentiality in terms of the manifestation relation.

The choice between the two kinds of chain is difficult to settle and will be discussed more deeply across the chapter. Interestingly, Vetter formulates (COULD) saying that a potentiality is an earlier stage *in* or *within* an iterated potentiality, not *of* an iterated potentiality, which could suggest that the iterated potentiality itself has stages, and so imply that an earlier stage is indeed part of the chain of repetitions within this potentiality, rather than being a potentiality that is external to it. However, this formulation cannot be taken as a definitive proof, because, in truth, she does not really express an opinion about what could be the nature of an 'earlier stage', so we can only speculate.

In any case, whichever of the two views we choose, the same problem remains: how are these potentialities connected to each other? If an earlier stage is part of the chain of repetitions within the same iterated potentiality, then how are the potentialities that are part of this chains connected to each other? If an earlier stage is a separate potentiality within a chain of yielding, then how do we explain a 'fewer times'-iterated potentiality yielding a 'more times'-iterated one?

In what follows, I will try not to commit entirely to one or the other view, presenting how both could work. This is because Vetter herself seems to swing between them. In general, her picture of iterated potentiality is quite confusing on this point, so that the nature of the 'earlier stage' relation is open to various interpretations. The idea that an earlier stage in an iterated potentiality is somehow part of the chain of repetitions is supported by Vetter's talking about stages that are *within* an iterated potentiality. However, the fact that these stages are taken by Vetter to correspond to antecedents and consequents of counterfactuals could suggest that some should be yielding others, rather than just be part of a series of repetitions.

The idea that iterated potentiality comes in stages seems quite a crucial aspect of this kind of potentiality, and very much essential to an understanding of it that is not purely intuitive but properly metaphysical. Since Vetter does not clarify what it means, metaphysically speaking, for an iterated potentiality to have stages, in what follows, we will try to offer some suggestions about how to interpret this idea. Some of the things we will discuss could be easily applied to both kinds of chains we have just presented, while others can only apply if we favour one over the other.

2. The Process of Iteration

Even if the concept of iteration is clearly metaphysical in Vetter's account, the idea of 'being an earlier stage' is difficult to clarify. As we said, looking at what Vetter has to say about iterated potentiality, she does not really explain what she means by 'stages', so that we do not have an idea of the nature of them. However, we can still speculate on what options we have in trying to explain what these 'stages' are and how they are connected.

Let us consider an example: "If I learned to play piano, I could become a pianist for the Royal Opera House". At the moment, I cannot play piano, and so even if I had the opportunity, I could not become a pianist for the Royal Opera House. However, I do have a once-iterated

potentiality to learn to play the piano, and so I have a twice-iterated potentiality to play the piano, which means that I also have a thrice-iterated potentiality to become a pianist for the Royal Opera House (if my learning goes incredibly well).

Considering that, by (COULD), our example is true if the potentiality to learn to play the piano is an earlier stage in a potentiality to become a pianist for the Royal Opera House, then it seems quite clear that the 'stages' in this example should be potentialities, so that the 'first stage' is the once-iterated potentiality, the 'second stage' is the twice-iterated potentiality, and the 'third stage' is the thrice-iterated potentiality, and so on.

It seems reasonable to think that these three potentialities are somehow connected to each other, because they appear to be in a sequence: the thrice-iterated potentiality seems to follow the twice-iterated potentiality which in turn seems to follow the once-iterated one. It looks like there is some sort of *process* going first from the once-iterated potentiality to the twice-iterated one and then from the twice-iterated potentiality to the thrice-iterated one, which we could call '*process of iteration*'. Let us see how this process could be explained in terms of the two kinds of chain we presented above.

In terms of 'chain of repetitions', it seems clear that these three potentialities should be connected by the fact that they represent different repetitions of the 'have a potentiality' function in one and the same sequence that constitute an iterated potentiality, and so that the process of iteration should be a kind of process of repetition.

Therefore, we could get the following picture of the chain connecting the three potentialities above and of the process of iteration: the once-iterated potentiality to learn to play the piano is the first repetition of the clause 'have a potentiality to' in 'have a potentiality to (1) have a potentiality to (2) play the piano' (the twice-iterated potentiality) and in 'have a potentiality to (1) have a potentiality to (2) have a potentiality to (3) become a pianist for the Royal Opera House' (the thrice-iterated potentiality), while the twice-iterated potentiality to play the piano is, in turn, the second repetition of the same clause in 'have a potentiality to (1) have a potentiality to (2) have a potentiality to (3) become a pianist for the Royal Opera House' (the thrice-iterated potentiality). In this way, the three potentialities are connected by being part of the series of repetitions that construct the thrice-iterated potentiality to become a pianist for the Royal Opera House.

In terms of 'chain of yielding', instead, with 'process of iteration' we mean the process by which 'fewer times'-iterated potentialities yield 'more times'-iterated potentialities. Even if we abandon the idea that these three potentialities are part of a chain of repetitions, still we seem justified in thinking that these three potentialities form a chain and that the process of iteration so intended can still connect several potentialities in sequence.

Therefore, we could get the following picture of the chain connecting the three potentialities above and of the process of iteration: a potentiality to learn to play the piano *yields* a potentiality to have a potentiality to play the piano which *yields* a potentiality to have a potentiality to have a potentiality to become a pianist for the Royal Opera House. In this case, the three potentialities are not part of the same series of repetitions but are still connected to each other by the fact that they give rise to each other in sequence.

In both cases, the fact that I possess a once-iterated potentiality to learn to play the piano allows me to possess a twice-iterated potentiality to play the piano, which in turn allows me to possess the thrice-iterated potentiality to become a pianist for the Royal Opera House. This means that the thrice-iterated potentiality belongs to me *because* I have the other two. In the case of the chain of repetitions, this is because the two other potentialities are the constituents of the series of repetitions that construct the thrice-iterated potentiality itself. In the case of the chain of yielding, this is because the possession of the thrice-iterated potentiality depends somehow on the possession of the other two. In any case, given our rough interpretation of counterfactuals as expressing the idea that the consequent is an outcome of the antecedent, with these two interpretations of the process of iteration we can confirm that our counterfactual example is true.

Note that, following both interpretations, we do not need to assume that an '$n > 1$'-iterated potentiality is a potentiality of a different kind compared to any once-iterated potentiality, as we can think that it is just a potentiality like any other, but whose possession is the result of the possession of other potentialities. My thrice-iterated potentiality to become a pianist for the Royal Opera House may not be different in nature from my twice-iterated potentiality to play the piano, or my once-iterated potentiality to learn to play the piano, but it is a potentiality that I can possess *because* I can possess the other two.

However, even if the nature of iterated potentiality as an entity may not require extra metaphysical investigation, the fact that its possession is the result of the possession of other potentialities (either because these are the component of the series of repetitions within the

iterated potentiality or because they yield it) does. It is this connection that must be metaphysically investigated and clarified, because it is what really makes a potentiality an iterated potentiality.

For our purposes then, we do not need to investigate the nature of iterated potentiality as an entity, so we can leave that question open. It is another question that is more crucial for determining whether Vetter's account of counterfactuals is acceptable: how is iterated potentiality connected to other potentialities? If we agree that it is via a chain of repetitions, how does this work? If we agree that it is via a chain of yielding instead, what does it mean for the potentialities involved 'to *yield*' one another, or 'to *depend*' on one another? Our interest from now on will be in trying to answer these questions, by suggesting ways to explain how the possession of an iterated potentiality can result from the possession of the potentialities that come before in the sequence.

3. Asymmetry, Transitivity, Non-reflexivity

For our account of counterfactuals in terms of potentiality to work, it is essential to explain the process of iteration. Depending on whether we envisage this process as a chain of repetitions or a chain of yielding we can attribute different features to it. For example, if we support the idea that it is a chain of yielding, we are inclined to think that it should involve some form *dependence* of an iterated potentiality on its earlier stages.

However, intuitively, we can attribute some characteristics to this process of iteration that seem independent of which "chain" view we adopt, and so that we could take to be common to both of them: asymmetry, transitivity, and non-reflexivity. Therefore, in what follows, we will use the term 'to iterate' as the general verb that represents the connection between potentialities and iterated potentialities within the process of iteration, which can be interpreted differently whether we adopt the "repetition" view or the "yielding" view. In this way, the expression 'potentiality *a* iterates iterated potentiality *b*' can mean either 'potentiality *a* is a previous repetition in the series of repetitions within iterated potentiality *b*' or 'potentiality *a* yields iterated potentiality *b*', depending on which view we prefer. Also, across this chapter we will keep talking about "'earlier stage' relation" and "iteration relation" as synonyms, standing for the relation that stands between the potentialities that are part of the

process of iteration in general, before deciding whether it should be expressed in terms of 'repetitions' or 'yielding'.

Asymmetry. It seems clear that the process of iteration tends to be one-directional: it goes from earlier stages in the sequence of potentialities to later stages in the same sequence. If some water has a once-iterated potentiality to freeze (and so turn into ice), and a twice-iterated potentiality to break (because then the ice can be broken), it seems obvious that the twice-iterated potentiality for this water to break can only follow and not precede the once-iterated potentiality for this same water to freeze. Therefore, the 'earlier stage' relation seems to be a unilateral relation, going only in the direction from an earlier-stage potentiality to a later-stage one, so that the possession of a later-stage potentiality is the result the possession of an earlier-stage one, but not vice versa. Even if it seems that there could be situations in which the process of iteration can work in both directions, for example, if we think that the potentiality to be Hesperus iterates the potentiality to be Phosphorus and vice versa, this is not the case. Indeed, as we said in previous chapters when talking about the necessity of identity, the potentiality to be Hesperus and the potentiality to be Phosphorus are simply the same potentiality, and so they are not two potentialities that can be different stages in the process. Also, they are equally iterated (i.e., they are both once-iterated potentialities) while the process should go from 'fewer times'-iterated potentialities to 'more times'-iterated potentialities. Therefore, in general, it seems correct to assume that, if the potentiality to be *F* iterates the potentiality to be *G*, then it is not the case that that the potentiality to be *G* iterates the potentiality to be *F*. Which means that the 'earlier stage' relation is *asymmetric*: when a potentiality to be *F* is an earlier stage in a potentiality to be *G*, then it is not the case that a potentiality to be *G* is an earlier stage in a potentiality to be *F*.

Transitivity. We said that there can be multiple stages in a process of iteration, and it seems to be the case that every later stage comes from the same earlier stages in the same sequence of potentialities. A cloud of steam has a once-iterated potentiality to be at a liquid status (turn into water), a twice-iterated potentiality to freeze (with this water turning into ice), and a thrice-iterated potentiality to break (because this ice can be broken). If the once-iterated potentiality to be at a liquid status iterates the twice-iterated potentiality to freeze, and the twice-iterated potentiality to freeze iterates the thrice-iterated potentiality to break, then the once-iterated potentiality to be at a liquid status iterates the thrice-iterated potentiality to break as well. Therefore, the 'earlier stage' relation seems very clearly *transitive*. If a potentiality to be F is an earlier stage in a potentiality to be *G*, and a potentiality to be *H* is an

earlier stage in a potentiality to be *F*, then a potentiality to be *H* is an earlier stage in potentiality to be *G*.

Non-reflexivity. Considering our assumption that the process of iteration involves several separate potentialities from 'fewer times'-iterated potentialities to 'more times'-iterated ones, it seems also the case that a potentiality should not iterate itself, because the process should involve potentialities that are "numerically differently" iterated (a once-iterated potentiality should not iterate other once-iterated potentialities). Naturally this is debatable, because trivial counterfactuals seem to involve exactly a 'self-iterating' potentiality: a counterfactual like "If I were Italian, I would be Italian" does not seem to involve more than one potentiality, which apparently iterates itself and remains once-iterated. However, we already said that in this instance we are only interested in "regular" counterfactuals, so we will not tackle this issue here. Therefore, we can assume, for simplicity, that the 'earlier stage' relation is *non-reflexive*.[43]

3.1. The Problem of Transitivity

These three aspects of the 'earlier stage' relation seem intuitively right, but one of them can lead to a strong case against Vetter's semantics of counterfactuals.[44] As Vetter herself says, there can be an unlimited number of iterations between two potentialities, because of the transitivity of the 'earlier stage' relation. This transitivity could turn into an issue for Vetter's account of counterfactuals because, as it stands, this account seems to enforce the truth of counterfactuals regardless of the number of iterations between the two potentialities involved. In (COULD), the only requirement for a counterfactual ascribing potentiality *a* in the antecedent and potentiality *b* in the consequent to be true is that potentiality *a* should be an earlier stage

[43] It was pointed out to me that if the 'earlier stage' relation is transitive, then it can be reflexive. Consider the following: you have the potentiality to be sitting, then standing, then sitting again. So, if the process of iteration is transitive, it should follow that the potentiality to sit can iterate the potentiality to sit. My answer to that could be that here the two potentialities to be sitting are not exactly the same: the first is a once-iterated potentiality to be sitting, while the second is a thrice-iterated potentiality to be sitting (being a potentiality following a twice-iterated potentiality to be standing). However, I admit that this issue would require a much longer investigation and discussion. In this work, I will accept the simplification above, but I do not take it to be final.

[44] I partly discussed this problem and its possible solution in my article in *Argumenta* (Casini, 2022).

of potentiality *b*, without any specification of "where" these two potentialities stand in the sequence.

Take a counterfactual that we would normally consider false, like:

(1) If this cloud of steam turned into water, it could break

By applying (COULD), we can see that this counterfactual should come out as true if the cloud of steam has an iterated potentiality to break (without any specification on how many times should be iterated), and a potentiality for this could of steam to turn into water is an earlier stage in this iterated potentiality (without any specification on how close the earlier stage should be to this iterated potentiality in the sequence). As we can see, in these conditions, (1) should be evaluated as true: on one side, from what we said, a cloud of steam has indeed a thrice-iterated potentiality to break; on the other, the potentiality to turn into water is indeed an earlier stage of the potentiality to break (because it is an earlier stage of the potentiality to turn into ice, that is an earlier stage of the potentiality to break).

Therefore, without any specification on the number of iterations allowed, concerning both how many times a potentiality should be iterated and how much "earlier" the earlier stage in the process should be, Vetter's account leads us to accept as true counterfactuals that we would normally consider false because the number of iterations implied is simply too high.

Once again Vetter's lack of clarity in defining the stages of iterated potentiality does not help. Metaphysically speaking, she does not impose a hierarchy over the stages, for example by saying that only the possession of the 'immediately-before' or 'directly-related' earlier-stage-potentiality can count towards the possession of a certain iterated potentiality. Nothing exclude that I can possess a ten-times-iterated potentiality simply by possessing the once-iterated potentiality in that sequence. Therefore, in terms of counterfactuals, it is hard to see how we can reject those involving two potentialities with an enormous number of iterations between them.

However, the problem arises from the specific analysis of counterfactuals coming from (COULD), rather than from the very idea of iterated potentialities. Metaphysically, it would be odd imposing some restriction on the transitivity of the 'earlier stage' relation, because any restriction would have to be an arbitrary cut-off point, which Vetter strongly rejects. As Vetter rejects that there should be cut-off points on the spectrum of potentialities going from minimal to maximal degree, so we should assume that she would not wish to impose an arbitrary

metaphysical limit on how far back the transitive relation goes between potentialities in a sequence. Vetter solved the problem of the spectrum of potentialities by appealing to the context, so that it is the context that sets the threshold on the potentiality degree. Maybe the context could be of help in this case too.

As we have seen, the 'earlier stage' relation cannot be left unrestricted when dealing with counterfactuals, because its transitivity would make us evaluate as true counterfactuals that we actually consider false. However, it is well-known (and Vetter herself agrees) that counterfactuals are context-sensitive. Therefore, a possible solution could be appealing to the context sensitivity of counterfactuals to avoid enforcing the truth of these counterfactuals. Counterfactuals are context-sensitive within the potentiality account in the sense that the potentialities that are considered when evaluating a counterfactual should be constrained by the context in which the counterfactual is asserted. Hence, the context could be the restriction required on the 'earlier stage' relation for the semantics of counterfactuals. Only the number of iterations that are contextually relevant for the counterfactual would count in the counterfactual evaluation. As we would not take "I can play the piano" to be true in a certain context (e.g. the pianist at a concert is ill and a substitute is required immediately) simply because I am a human and I can learn how to play the piano, despite me having indeed a twice-iterated potentiality to play the piano, so we would not take our example to be true simply because with enough iterations we can get from the potentiality for some cloud of steam to turn into water to the potentiality for some ice to break.

Therefore, we could suppose that, while metaphysically speaking there is indeed a process of iteration from the cloud of steam turning into water and its breaking, which are connected through some transitive steps of the 'earlier stage' relation, in terms of context the leap required over the stages of this process is too wide for it to be relevant to the truth conditions of the counterfactual.

To see how this could work consider (1) again:

(1) If this cloud of steam turned into water, it could break

If we restrict (COULD) in some ways, for example by saying that (1) is true in context *C* only if the cloud of steam has, in context *C*, say, a *once-iterated potentiality* to break or only if a potentiality to turn into water is, in context *C*, an *immediate earlier stage* of a potentiality to break (immediate as 'there are no other stages in between the two potentialities'), then we can easily get to our correct conclusion that (1) is false. It could be the context that determines how

many times the potentiality ascribed by the consequent can be iterated to be relevant in that context, or how many stages there can be between the two potentialities ascribed by the counterfactuals, for this to be true in that context.

However, such appeal to the context is problematic because even if we know that the context may determine that the leap is too wide, that still requires explaining. A context always needs a metaphysical background on which to operate and make a selection of what is relevant, so once again we need to have a clearer metaphysical explanation of these "stages" over which the leap happens.

I am not going to offer a solution to this problem here. It is a very serious problem for Vetter, and it could represent a strong case against her semantics of counterfactuals. The appeal to the context could help, but it cannot be effective until there is a clear metaphysical picture for it to work on. The selection that the context must make happens on the 'earlier stage' relation, and we still do not know what the nature of this relation is. Only once this would be clarified we could try to restrict it by using the context. However, since here I am only going to speculate on various possible scenarios, this line of work will not be pursued any further.

4. Modal Iteration

There is another important aspect of the 'earlier stage' relation in general that needs to be considered: for this relation to be effective for an account of counterfactuals it must have a modal connotation.

Indeed, the distinction between could- and would-counterfactual in terms of the consequent being *possible given the antecedent* or *necessary given the antecedent*, which was the key for a formulation of (WOULD) in the previous chapter, can be explained only if the connection between the potentialities involved in a counterfactual has *modal nature*. Following our rough interpretation of how counterfactuals work, it is the antecedent that *makes* the consequent *necessary or possible* within a counterfactual, so the 'earlier stage' relation should be able to map this semantic intuition. Therefore, the connection between the potentialities in antecedent and consequent should be characterized in a modal way or, at least, it must have modal consequences: the 'earlier stage' relation between these potentialities must imply some sort of necessity or possibility for the later-stage potentiality to be possessed if the earlier-stage potentiality is possessed.

In what follows, we will explore two ways in which this modal connotation can be implemented. First, we will consider the proposal we introduced in the previous chapter concerning existential quantification over chains of potentialities. Afterwards, we will show why the problem of this modal aspect cannot be solved by simply considering the 'earlier stage' relation a kind of modal relation.

4.1. Existence of Chains

When we applied the different stages of the formulation of (WOULD) to some examples of counterfactuals in the previous chapter, we showed that we need to include this extra modal notion for Vetter's account to offer some adequate truth conditions for those counterfactuals.[45] Only under this *modal* interpretation of the 'earlier stage' relation the examples could get some effective truth conditions that are quite in line with our intuitions.

Such an interpretation was that the relation of a potentiality being an earlier stage in another iterated potentiality is connected to the *existence* of a chain of potentialities including them both. To provide a modal interpretation of this relation, it was suggested that the *existence* of this chain could be taken to mean that an object's possession of the earlier potentiality *makes possible* the possession of the later iterated potentiality by the same object as well, while the *non-existence* of this chain could be taken to mean that an object's possession of the earlier potentiality *makes impossible* the possession of the later iterated potentiality by the same object. Note that the term 'chain' here can refer both to the chain of repetitions or the chain of yielding, depending on which approach we prefer. In general, when discussing the examples in the previous chapter, it was used purely in a metaphorical way, intending a sequence of potentialities somehow connected to each other.

Vetter's formulation already allowed us to interpret (COULD) as saying that for the truth of a could-counterfactual the consequent should be *possible given the antecedent*. However, it was the addition of the notions of '*making possible*' and its opposite '*making impossible*', concerning the possession of a potentiality in view of the possession of another, that allowed us to interpret (~COULD) as saying that, for the truth of a negated could-counterfactual, the

[45] Parts of this discussion of 'chains of potentialities' have previously appeared in my article (Casini, 2022)

consequent should be *impossible given the antecedent*, and (WOULD) as saying that, for the truth of a would-counterfactual, the consequent should be *necessary given the antecedent*. These are all extremely desirable and intuitive interpretations, so the '*making possible/impossible*' formulation connected to the existence or non-existence of chains of potentialities seemed a viable option to solve this issue.

However, the metaphor of a chain could only go so far, because again it was not explained what sort of relation there is between its "chain links". We tried not to deal with this by suggesting that what matters is the *existence* of this chain, rather than the relation connecting its links. The modal aspect required was attributed to an existential quantification over the chain rather than to any characteristic of the chain itself, so that this modal aspect would be there independently of whether it is a chain of repetitions or a chain or yielding. It seemed a sensible solution and loosely based on a familiar philosophical tradition which links existential quantification (over possible worlds) to possibility. Above, we worded this modal aspect in terms of '*making possible*' or '*making impossible*', which could be quite unclear, so let us rephrase the idea more clearly. On one side, the fact that there is a chain connecting two potentialities is taken to correspond to the fact that it is possible for an object to possess the second (iterated) potentiality if this object possesses the first potentiality. On the other, the fact that there is not a chain connecting them is taken to correspond to the fact that it is impossible for the object to possess the second potentiality if it possesses the first.[46]

Even if this view is comforting in having such a familiar look, and seems plausible enough, there are clearly some issues with it. One problem is that, at the end of the day, the idea of a chain of potentialities implemented then was only a metaphor, so we need to find a way to

[46] There is an interesting question that we could ask at this point: this connection between potentialities, that we described in terms of chains and of process of iteration, does it happen between general 'uninstantiated' potentialities or between the potentialities that are instantiated in the particular situation involved in a counterfactual? I.e., is it the "*universal*" potentiality to learn to play the piano that is an earlier stage in the "*universal*" potentiality to play the piano, so that, in any situation or circumstances, whenever an object has a potentiality to play the piano, a potentiality to learn to play piano will always be its earlier stage? Or is it the "*trope*" potentiality to learn to play the piano that is an earlier stage in the "*trope*" potentiality to play the piano, so that this connection depends on the specific situation in which the two universal potentialities are instantiated? Implicitly, I assumed across this work that this connection operates at the level of universals, and I will continue to assume so for the remainder of it, but this could easily turn into a matter for debate.

explain more precisely on what we would existentially quantify in this case: whether it is a chain of repetitions or a chain of yielding.

However, the main problem with this solution is that, even if we clarify the metaphor of the chain, we end up in a circle. If, for example, instead of talking about chains, we talk about a dependence relation in general, we could rephrase our view as follows and immediately see the issue. On one side, the fact that there is a dependence relation between the possessions of two potentialities is taken to correspond to the fact that it is possible for an object to possess the second (iterated) potentiality if this object possesses the first potentiality. On the other, the fact that there is not such a dependence relation between these possessions is taken to correspond to the fact that it is impossible for the object to possess the second potentiality if it possesses the first. As we can see, this solution just goes back in circle, because it looks like we are simply saying that, for the possession of a potentiality to depend on the possession of another, there must be a relation of dependence between their possessions: for it to be possible that an object possesses a potentiality depending on it possessing another, there must exist a dependence relation between the possession of the former potentiality and the possession of latter. A similar reasoning could be done if we consider this explanation in terms of a repetition relation: the fact that there is a series of repetitions that results in an iterated potentiality results in the possession of this potentiality. In general, what this "existential" view suggests is simply this: if there is a connection between the two potentialities, then they are connected; if there is not a connection between them, then they are not connected. This explanation is clearly circular so it cannot work.

The metaphor of a chain, then, even if it seems to work at an intuitive level, cannot be accepted once we attempt to discuss it into more details. Therefore, we need to find alternative ways to introduce a modal aspect in the 'earlier stage' relation.

4.2. Modal Relation

One option on how to deal with this modal aspect that is worth exploring is to consider whether the 'earlier stage' relation should simply be a pure modal relation: a relation so that, *necessarily*, if potentiality *a* were possessed, then potentiality *b* would be possessed.

We can see three evident reasons why this cannot be the case. First, taking a modal relation like this to be primitive would be quite a big blow to Vetter's objective of offering an

explanation of modality itself, because her whole attempt to reduce possibility and necessity to potentiality would just be nullified. Second, this relation is formulated as a would-counterfactual, and if it is itself a counterfactual, it cannot be used to explain other would-counterfactuals. Third, it leaves could-counterfactuals out of the picture. A modal relation so formulated is too strong to include could-counterfactuals. In looking at our application of (COULD) and (WOULD), it seems that the phrase 'being an earlier stage' must not imply the *necessity* for the later stage to be possessed, because in could-counterfactuals this notion must be used to express only the possibility for the consequent to be the case given the antecedent, so that we can derive the necessity of the consequent given the antecedent of would-counterfactuals via interdefinability (from 'being an earlier stage of *p*' to 'not being an earlier stage of non-*p*').

However, even if we tried to formulate a corresponding pure modal relation for could-counterfactuals in terms of possibility, we would still face the first and the second problem: taking it as primitive would jeopardise the primary aim of Vetter's project and formulating it as a could-counterfactual would make it useless to account for other could-counterfactuals. Therefore, for all these reasons, the 'earlier stage' relation cannot be considered simply a modal relation.

Given all this, the best option in trying to account for this modal connotation is to compare the 'earlier stage' relation with types of relations that can easily be given a modal reading. Indeed, the kinds of relation that we are going to discuss in the next section can easily be read in a modal way and so account for the modal flavour that the 'earlier stage' relation requires.

5. Manifestation, Causation, and Grounding

In the last two sections, we considered those aspects of the 'earlier stage' relation that are quite general and that are in common with both "chain" views: the fact that this relation is asymmetric, transitive, and non-reflexive, and the fact that this relation must present a modal interpretation. Now we will try to focus on the details of this relation and on how exactly we can characterize it as a chain of repetitions on one side, or as a chain of yielding on the other. In doing this, we will consider three kinds of relation which could explain the two kinds of chain. The first one, the manifestation relation, seems to work if we adopt the 'chain of repetitions'

view, while it is excluded if we adopt the 'chain of yielding' view. The other two, causation and grounding, are instead two possible dependence relations that could represent the idea of 'dependence' between the potentialities belonging to a chain of yielding. All three of them face some difficulties and, as said, I do not aim to settle the issue about which view is the best here, but I only want to speculate on some alternative solutions and their problems.

5.1. Repetition and Manifestation

Let us focus first on the chain of repetitions view. This is the view that an earlier-stage potentiality is simply one of the repetitions of 'have a potentiality to' in the sequence 'have a potentiality to (1) have a potentiality to (2) … have a potentiality to (*n*) something' in an '*n* times'-iterated potentiality. In this case, the connection between the potentialities is due to the fact that they are part of the same series of repetitions of the 'have a potentiality' function that concludes with the iterated potentiality in which we are interested.

A connection of this sort is extremely general because Vetter is extremely liberal concerning iterated potentialities, in a way that allows this connection to be almost limitless. Given that there could be infinite repetitions of the 'have a potentiality' function in an iterated potentiality, and that these repetitions can connect essentially any potentialities, applying this relation to counterfactuals means that the scope of an account of counterfactuals in these terms is equally limitless. In a way, the 'earlier stage' relation so conceived do not need to be limited to any kind of relation, but it can stand for any possible sort of connection that there could be between an antecedent and a consequent, whether this is a logical connection, a dependence connection or else. This means that any kind of counterfactual can be included in this account, without any limitation concerning the idea of a dependence between antecedent and consequent, and this is clearly an advantage of this view.

However, a relation so extremely unspecified necessarily leaves some essential details unexplained. For example, how is it exactly that a potentiality to learn to play the piano corresponds to the first repetition of 'have a potentiality to' in 'have a potentiality to have a potentiality to play the piano'?

My main issue with this view is with this correspondence: how do we know that the first repetition corresponds to that potentiality exactly? Could it correspond to other potentialities as well, like the potentiality to have a piano in the room or the potentiality to buy a piano? If we

take the two counterfactuals "If I learned how to play the piano, I could play Mozart's Piano Concerto No 24" and "If I bought a piano, I could play Mozart's Piano Concerto No 24", it seems that the earlier-stage potentiality in the twice-iterated potentiality to play Mozart's Piano Concerto No 24 is different, so that the repetitions in the chain can actually corresponds to different potentialities. In one case, the first repetition of 'have a potentiality to' corresponds to 'have a potentiality to learn to play the piano'; in the other it corresponds to 'have a potentiality to buy a piano'. Could this mean that any potentiality could correspond to this first repetition? Is there anything that blocks us from saying that 'have a potentiality to' in this case could correspond to something like 'have a potentiality to eat an apple'?

As you can see, my worry with this is multifaceted. On one side, this solution does not seem to provide enough content for the antecedent to be exactly and precisely identified. The idea that an earlier-stage potentiality corresponds simply to a repetition of 'have a potentiality to' is basically too generic, because 'have a potentiality to' does not provide any real content for this potentiality in a way that can fix the antecedent of a counterfactual. On the other side, and in many ways consequently, this solution does not seem to solve at all the problem of explaining the connection between antecedent and consequent, because if any potentiality can be the earlier stage of another potentiality, then any pair of antecedent and consequent can produce a true counterfactual, independently of any real connection between them. This last worry is probably extreme, and possibly unfair, because Vetter does not seem to support a complete lack of 'coherence' between the stages in a chain of potentialities that is connected to a counterfactual, even if it is a chain of repetitions. This 'coherence' might once again come from an appeal to the context, as in the case of a restriction over transitivity seen in section 3.1., so that only earlier-stage potentialities that are the most relevant for the construction of the iterated potentiality counts towards the truth of a counterfactual.

However, even if this last worry can be tackled relatively easily, my other two worries remain: how does the correspondence between repetitions and content-specific potentialities work? How can a repetition of this sort provide a real content for the antecedent?

Also, can we really leave the relation between the repetitions in an iterated potentiality unexplained? The 'chain of repetition' view rejects that this should be explained in terms of dependence, so that it denies that the 'earlier stage' relation should hold only between potentialities that depends somehow on one another. This rejection could be an advantage, as we know, because it allows the account to cover many more counterfactuals than those which

clearly express a dependence of the consequent on the antecedent. However, the problem then is how we can characterize the connection between the repetitions of the 'have a potentiality' function in a way such that it can represent a satisfactory background to counterfactuals. The simple idea of a 'repetition relation' does not seem able to map the connection between antecedent and consequent in a way that makes it justice. My support for the presence of a dependence relation at least in the standard cases of counterfactuals comes from my belief that our understanding of counterfactuals is primarily guided by the fact that antecedent and consequent are connected, and that this connection has a substantial value in determining the meaning of a counterfactual. Therefore, I cannot help finding the simple idea of a relation of 'repetition' too weak for this purpose.

However, maybe a 'repetition relation' does not need to be taken as the true relation behind the 'earlier stage' relation, because it could actually correspond to another relation. Vetter herself has given us some clue on what this other relation could be. One of the most important relations in her framework is the manifestation relation, which is the relation between a potentiality and its manifestation and is taken to be primitive. Therefore, we could wonder whether the 'earlier stage' relation is, in fact, nothing more than the manifestation relation itself, where the manifestations of the earlier-stage potentialities precede the manifestations of the later-stage potentialities.

Take the once-iterated potentiality to learn to play the piano and the twice-iterated potentiality to play the piano. It seems plausible to think that the potentiality to learn to play the piano manifests first, because it manifests in 'learning to play the piano', while the iterated potentiality to play the piano manifests afterwards, in 'playing the piano', so that the manifestation of the former is 'earlier' than the manifestation of the latter. We can make this interpretation in terms of 'earlier' and 'later' if we assume that potentiality is forward-looking in time, but in general it seems the case that these manifestations happen in sequence. In this picture, a potentiality to learn to play the piano has a manifestation, 'learning to play the piano', that leads to a potentiality to play the piano, because the manifestation 'playing the piano' follows 'learning to play the piano'. Therefore, a chain of potentialities could be taken to correspond to a series of potentialities whose manifestations happen in sequence, concluding with the manifestation of the relevant iterated potentiality.

If we accept this view, we can define an earlier-stage potentiality as a potentiality whose manifestation is part of the sequence of manifestations that ultimately leads to the

manifestation of the iterated potentiality. This view has the advantage to be relatively simple and to rely on a relation, the manifestation relation, which is "stronger" than an unidentified 'repetition relation' but which is still assumed to be primitive by Vetter, and so does not need some kind of explanation in terms of dependence. Also, it could easily account for the modal aspect of the 'earlier stage' relation that we discussed above: if the existence of a potentiality with *p* as its manifestation makes *p* possible, then it might be that a potentiality whose manifestation makes possible the manifestation of another potentiality makes the latter potentiality's possession possible, so that the earlier-stage potentiality, by having a manifestation that makes possible the manifestation of the iterated potentiality, makes possible the possession of the iterated potentiality.

However, there are some difficulties with this account. The fact that the manifestation relation is primitive should mean that we can leave unexplained what kind of relation there is between the so-called 'stages'. Assume that we accept the manifestation relation as the only metaphysically relevant relation between an iterated potentiality and the potentialities in the sequence. The 'stages' in an iterated potentiality then are just potentialities whose manifestations precede the manifestation of the iterated potentiality itself. Therefore, a counterfactual "If *x* were *F*, then *x* could be *G*" is true if a potentiality for *x* to be *F* has a manifestation that precedes the manifestation of the iterated potentiality for *x* to be *G*, and this should be the only explanation we need.

This view is clearly unsatisfactory in terms of a chain of yielding, because this explanation does not seem enough to account for the connection between antecedent and consequent so that it is *because x* is *F* that then *x* can or must be *G*. Such a picture of the 'earlier-stage' relation in terms of manifestation does not really explain how a 'fewer times'-iterated potentiality can relate to a 'more times'-iterated one in a way that is connected to the truth of a counterfactual, if this is intended in terms of a dependence between antecedent and consequent. In other words, the manifestation relation does not seem to clarify, for example, how is that my potentiality to learn the piano relates to my iterated potentiality to play Mozart's Piano Concerto No 24 so as to ensure that "If I were to learn to play the piano, I could play Mozart's Piano Concerto No 24" is true, because it does not explain how the consequent can depend on the antecedent.

Therefore, an explanation in terms of manifestation clearly does not work for the "yielding" view. It works better for the "repetition" view because this view does not rely on a

definition of the 'earlier stage' relation based on dependence, but on the fact that the repetitions ultimately constitute the iterated potentiality. Therefore, we could explain the relation between the repetitions in terms of the manifestation relation, suggesting that each repetition in an iterated potentiality is connected to the other by the fact that its manifestation precedes the manifestation of the other, in a sequence of manifestations that concludes in the iterated potentiality itself. This interpretation feels vary natural and the fact that the "repetition" view allows for it is another point in its favour.

However, I wonder if the same problems I had with the 'repetition relation' extend to the manifestation relation as well, since it is a primitive relation. Can we really leave the manifestation relation unexplained if we want to use it to account for counterfactuals? In general, even if we reject that it should be a dependence relation, still the 'earlier stage' relation employed by Vetter really does not seem intuitively clear enough to serve as the unexplained basis of a theory of counterfactuals, so that to link it to a primitive, unexplained, relation such as the manifestation relation seems to me that it would just leave us again with too many questions unanswered. Indeed, the application of the manifestation relation does not seem to help with my other worries concerning the definition of the content of the antecedent and the unclarity of the correspondence between repetitions of 'have a potentiality to' and content-specific potentialities.

In conclusion, even if the idea that the 'earlier-stage' relation should be reduced to the manifestation relation has its appeal for the 'chain of repetitions' view, still it seems to me that the relation between the two potentialities ascribed by a counterfactual should be a non-primitive one if we want a clear definition of the antecedent and of its connection to the consequent. Therefore, it is worth investigating the alternative proposals concerning the chain of yielding.

5.2. Yielding and Causation

After discussing what could happen if we adopt the 'chain of repetition' view, let us consider what could be the consequences of adopting the 'chain of yielding' view. In this case, we do not take the potentialities involved in the process of iteration as part of a series of repetitions that constitute an iterated potentiality, but we are considering them separately, as one giving rise or 'yielding' the other in a sequence until one of these potentialities yields the

iterated potentiality. In contrast with the 'chain of repetition' view, here the relation between the potentialities is less general and cannot really stand for any possible kind of relation that there can be between antecedent and consequent, because it relies on the idea of a dependence between them. However, being less general, and being between separate potentialities, this relation seems much more able to provide a content of the antecedent (which is simply the content of the separate earlier-stage potentiality) and avoids completely the problem of the correspondence between repetitions and content-specific potentialities, because the repetitions in this case do not correspond to any content but are just repetitions within separate potentialities. Also, by appealing to the idea of 'yielding' and 'dependence', this view cannot afford to leave the connection between potentialities unexplained and take it as primitive, which avoids the issue of this relation being too generic. However, this means that the supporters of this view must find a clear definition of this 'dependence relation'.

Therefore, the main issue with this view is to explain what is meant by 'yielding' and 'dependence' between potentialities, because this 'dependence' relation needs to be explained for the view to work. In what follows, we will consider two possible explanations, that connect this relation to two other 'dependence' relation: causation and grounding.

We start by considering whether this relation of dependence between potentialities should be one of the most common dependence relations, a *causal* dependence relation. In this way, an earlier-stage potentiality would be the *cause* of an iterated potentiality, which would be the *effect*. Describing this 'dependence' relation in these terms has the advantage of relying on a long philosophical tradition concerning causation. Also, by saying that the possession of an earlier stage is only a necessary but not sufficient cause for the possession of an iterated potentiality could help accounting for the modal aspect we need, because it would make this possession possible but not necessary given the earlier stage.

However, considering that this 'dependence' relation should only correspond to causal dependence seem somehow reductive, especially if this is used as a metaphysical background for counterfactuals. If this dependence were nothing more than causal dependence, then the scope of cases that this relation could cover would be quite limited and, in terms of an account of counterfactuals, it could offer an explanation for only a very restricted number of counterfactuals, those in which the relation between antecedent and consequent is clearly causal. For example, consider the relationship between the potentiality to be male and the potentiality to be human, or some other case in which the connection seems to involve logical

entailment rather than causal connection. In general, it seems that Vetter should take the connection between an iterated potentiality and its earlier stages as being much more general than a causal relation, so to cover a wider range of cases, and to make it acceptable for an account of counterfactuals. If the manifestation relation could be considered general enough in this sense but not enough explanatory, the causation relation instead could be taken to be explanatory enough, but not enough general, and so too restricting to produce a sufficiently wide account of counterfactuals. Therefore, we need to investigate another kind of dependence relation.

5.3. Yielding and Grounding

We have seen that trying to identify the 'earlier stage' relation with other kinds of relations is not easy, because we can find problems with each of them. After considering whether this relation is a modal relation, a manifestation relation or a causal relation, the last proposal that we are going to discuss here is whether the 'earlier stage' relation is a form of grounding. Indeed, if we adopt the "yielding" view, the fact that the possession of a potentiality is dependent on the possession of another could be a reason to think that this dependence relation is similar to grounding. Are we justified in making such a comparison?

Some justification could be found in what Vetter says about grounding. Vetter assumes that relations like grounding and 'because' are metaphysically real (see 2015, p. 26) and explains grounding as follows: some facts are *grounded in* other facts when these facts hold *in virtue of* or *because of* the other facts' holding. Therefore, we could re-formulate the process of iteration and the 'earlier stage' relation in terms of grounding as follow: potentiality *a* yields potentiality *b*, so that the possession of *b* depends on the possession of *a*, if the fact that a potentiality *b* is possessed *is grounded* in the fact that potentiality *a* is possessed, so that the fact that potentiality *b* is possessed holds *because of* the holding of the fact that *a* is possessed.

This seems a good way to explain the dependence between potentialities *a* and *b*, and it could even account for the modal aspect of this dependence: in grounding the possession of the iterated potentiality, it is reasonable to think that the possession of the earlier stage potentiality makes possible this possession of the iterated one.

However, once again there could be an issue in this proposal, if we consider that Vetter also assumes that grounding is related to fundamentality: if there is a fundamental level, then

its fundamentality could be characterized in terms of grounding (2015, pp. 26-27). For Vetter, comparative fundamentality too can be thought in terms of grounding: that which grounds is more fundamental that that which is grounded in it. Therefore, this might lead us to think that when the possession of potentiality *b* is grounded in the possession of potentiality *a* this means that potentiality *a* is more fundamental than *b*. In turn, this should imply that the more iterated is a potentiality, the less fundamental it is. Potentialities are for Vetter at the most fundamental level because, as she says, "as we progress from the less to the more fundamental levels, we will always find potentialities" (2015, p. 25), but it looks like that the process of iteration, if it is based on grounding, should go from more fundamental to lesser and lesser fundamental potentialities. My twice-iterated potentiality to play the piano should then be less fundamental than my once-iterated potentiality to learn to play the piano, so that we could wonder how fundamental a '100 times'-iterated or '1000 times'-iterated potentiality could be.

Also, this idea could turn quite dangerous for an account of counterfactuals that takes the 'earlier stage' relation as its metaphysical basis. Let us assume that this relation is grounding, and that grounding is fundamentality, as described above. By (COULD), the potentiality ascribed by the antecedent in a counterfactual is always said to be an earlier stage in the potentiality ascribed by the consequent. However, if 'being an earlier stage' means grounding, then we should say that the potentiality ascribed by the antecedent always grounds the potentiality ascribed by the consequent. But again, if grounding is fundamentality, this compels us to say that, for every counterfactual, the potentiality ascribed by the antecedent is *always* more fundamental than the potentiality ascribed by the consequent.

This is obviously not true, because there are true counterfactuals where the antecedent is not more basic than the consequent, like "If a certain coin had a certain shape, then it could be a triangle" or "If I were to be unmarried, I would be a bachelor". In general, the potentiality ascribed by the antecedent does not need to be more fundamental for a counterfactual to be true.

However, Vetter could respond by pointing out that her conception of grounding is *operational* rather than *relational* (see 2015, pp. 27-28). While relational grounding treats grounding as a relation between entities that is expressed by a two-place predicate operator '*x* grounds *y*', operational grounding takes grounding to be expressed by a two-place sentence operator completed not by terms but by sentences, like 'because' or 'in virtue of'. This means that Vetter is concerned with the dependence of an object's *having* some potentiality on other

objects' *having* other potentialities. In other terms, it is not so much that potentiality *a* grounds potentiality *b*, rather it is the fact that *a* is possessed that grounds the fact that *b* is possessed. This means that it is the possession that is more or less fundamental, not the potentiality itself.

This idea seems to work better with our account of counterfactuals. After all, we assume that counterfactuals ascribe the *possession* of potentialities to certain objects: a counterfactual essentially claims that an object is allowed or compelled to possess a certain iterated potentiality if it possesses another potentiality. However, the idea that the possession of the potentiality ascribed by the antecedent is always more fundamental than the possession of the potentiality ascribed by the consequent, despite not attributing more or less fundamentality to the potentialities themselves, still seems to suggest that the antecedent of a counterfactual should be more fundamental than its consequent. Even if it is not the ascribed potentiality to be more fundamental, still Vetter seems to say that its ascribed possession is, and therefore that in any case what is ascribed by the antecedent is more fundamental than what is ascribed by the consequent. And, for the truth of a counterfactual, this is simply not the case: a counterfactual can be true even if its antecedent is not more basic or fundamental than its consequent.

Therefore, it seems that we could consider the 'earlier stage' relation a form of grounding only if Vetter is able to amend her position on grounding and fundamentality. If we could negotiate this, then grounding might turn to be a good way to represent the 'earlier stage' relation. In truth, this view that this relation is a form of grounding can explain the idea of 'iteration' itself. In general, 'to iterate' means to repeat a process or to have something performing or running again, hence our alternative representation in terms of a chain of repetitions, where the 'stages' in an iterated potentiality are repetitions of the application of the 'have a potentiality' function: have a potentiality to have a potentiality to have a potentiality. In terms of grounding, we could interpret this by saying that it is the 'possession-fact' (the fact that a potentiality is possessed) that is repeated in an iterated potentiality, so iteration could be characterized in terms of one possession-fact grounding another possession-fact. If the repetitions in an iterated potentiality are repetitions of 'possession-facts' grounding one another in sequence, and if a potentiality yields another when the fact that the latter potentiality is possessed grounds the fact that the former potentiality is possessed, then, in this way, the chain of yielding could somehow merge with the chain of repetitions and solve our initial problem of how to define the stages.

Would this solve the issue of the generality of the 'earlier stage' relation? Is grounding general enough to allow to cover the same wide range of cases that the 'chain of repetition' view covers? I suspect that the answer depends on what one think 'grounding' is, which is very much a matter of debate. I do not wish to enter in that discussion here, so I will leave this question open.

In any case, though, until we settle the problem surrounding fundamentality, even grounding cannot be completely considered a good candidate for explaining the 'earlier stage' relation that we need for an account of counterfactuals.

6. Conclusion of Chapter VI

In this chapter we have investigated the relationship between counterfactuals and iterated potentiality. After explaining why Vetter suggested such a connection, we characterized the metaphysical background of counterfactuals in terms of two kinds of chains, a chain of repetitions and a chain of yielding (section 1) both of which can describe the so-called 'process of iteration' (section 2). We saw that the 'earlier stage' relation has some general features: asymmetry, transitivity, and non-reflexivity (section 3). Also, we explained that this relation should have a modal connotation to account for counterfactuals (section 4). Finally, we tried to explain this aspect by linking it to different kinds of relations, depending on which kind of chain we choose, including manifestation, causation, and grounding (section 5).

We have seen that, despite being essential for the formulation of effective truth conditions for counterfactuals, the connection existing between the potentialities involved in the process of iteration remains largely unexplained in Vetter's picture. We characterized this connection in terms of either a chain of repetitions, which involves an undefined and primitive relation, or a chain of yielding, which involves some kind of dependence relation that needs explaining. In investigating what sort of relation the 'earlier stage' relation could be, we considered different options, but none of them seemed satisfactory. Naturally, our investigation was limited to those proposals that seemed to arise most naturally from Vetter's work and understanding of potentiality.

In truth, Vetter seems to link the relation between an iterated potentiality and its stages to a lot of things. It does look like that there may just be fundamentally different kinds of cases that Vetter seems to be bunching together as involving 'iterated potentialities' – e.g., causal

cases, logical entailment cases, and perhaps others – so that it might be that there is not really any single relationship of 'iteration' that is fit to serve as the basis of her account of counterfactuals.

Given this, maybe we should conclude that iteration is not a form of dependence and abandon the 'chain of yielding' view. However, if iteration is not dependence, then what is it? Also, if it is not dependence, how can iterated potentiality be the metaphysical background of counterfactuals?

Maybe an extremely general definition of this relation is the only thing we can aim for, if we want the 'earlier stage' relation to cover all the cases that Vetter seems to understand under the label of 'iterated potentiality'. If this is the case, then the idea of a chain of repetitions seems the way to go. However, can such a general definition be enough to offer a clear metaphysical picture of iterated potentiality in relation to counterfactuals? My believe is that it cannot, because there are questions about the content of the antecedent and the correspondence of repetitions and content-specific potentialities that remain unanswered. Also, we seem to need quite an exhaustive picture to be able to apply any contextual restrictions, and these are crucial if we want to avoid the problems arising from the transitivity of the 'earlier stage' relation.

Maybe we should just take 'being an earlier stage' to mean a range of different things, with different sorts of relations being suitable for different sorts of cases (sometimes causal, sometimes logical, etc.). However, a relation conceived in these terms seems more based on an *ad hoc* attempt to accommodate different counterfactuals than on a truly metaphysical picture of potentiality. Accepting that the 'iteration' relation corresponds to different relations depending more on the semantic results we are looking for than on the metaphysical picture in which is inserted would make Vetter's account of counterfactuals simply a semantics-based account and not a metaphysics-based one. And an account purely based on semantic results is not what Vetter wants because her intents are primarily metaphysical.

Maybe the real problem is that iterated potentiality is not only one thing as Vetter suggests. Iterated potentiality is defined initially as potentiality to have a potentiality to have a potentiality etc…, which is simply the repetition of the 'have a potentiality' function, like me having a potentiality to have a potentiality to play piano. Hence the 'chain of repetition' view. However, afterwards, and for the most part, iterated potentiality seems to turn into a kind of potentiality that is somehow dependent on other potentialities, like my potentiality to play the

piano that is twice iterated because it depends on my potentiality to learn to play the piano. Hence the 'chain of yielding' view.

Let us consider the following picture. We have iterated potentialities, like the potentiality to have the potentiality to be *G*. The question is, though, how are we to explain how objects have these iterated potentialities? It seems that they arise from the possession of simpler ('fewer times'-iterated) potentialities: e.g., my iterated potentiality to have the potentiality to be *G* is owed to my potentiality to be *F*. But how exactly does that work? And similarly, how are those sorts of facts reflected in counterfactuals? Indeed, it seems that, in a case like the preceding one, sometimes the relationship between these potentialities ensures that if I were to be *F* then it would be possible for me to be *G*. But how does the relationship between the simpler potentiality to be *F* and the iterated potentiality to have the potentiality to be *G* ensures the truth of that counterfactual?

Is it really the case that for something to have a potentiality to have a potentiality to *p* is the same to have a potentiality to *q* that somehow 'causes' to have a potentiality to *p*? In ordinary language, to say that 'I have a potentiality to have a potentiality to play the piano' does not seem to imply that 'I have a potentiality to play the piano since I have a potentiality to learn to play the piano', because the expression 'have a potentiality to have a potentiality' does not make it clear precisely which once-iterated potentiality provides the basis for the twice-iterated one. How this correspondence works is left completely unexplained.

Maybe this is why Vetter's picture of iterated potentiality faces challenges, because it relies on a controversial premise: that iterated potentiality is taken to be one thing, while it is not. On one side, iterated potentiality is considered very simply to be the result of a repetition, without real specification of these repetitions, while on the other this potentiality is taken to be connected in a much more substantial way to others. Iterated potentiality seems to be many things, and to play many roles; roles that it may not be possible for just one thing to play.

Our conclusion is that significant questions remain about the nature of the 'earlier stage' relation and that, in particular, it remains to be shown that there is a satisfactory account of the metaphysics of iterated potentialities that can be used to provide an adequate account of counterfactuals. However, it does also remain to be shown that *there is not* a suitably satisfactory account of the metaphysics of iterated potentialities in relation to counterfactuals. Therefore, it may yet be that, with more work and refinement, Vetter's framework can be used to develop a correct account of counterfactuals.

CONCLUSION

Across this thesis we have discussed a new approach to counterfactuals, based on Barbara Vetter's potentiality account of modality.

After introducing counterfactuals in Chapter I, we showed that they should be classed as part of modal discourse in Chapter II. However, in Chapter III, we suggested that the kind of modal discourse in which they should be inserted does not need to be the traditional possible world account of modality, but it could be instead the theory of possibility put forward by Vetter, potentialism. We discussed the details of how counterfactuals relate to potentiality in Chapter IV, by discussing both the metaphysical background and the semantic framework surrounding this account of counterfactuals. In Chapter V, we tried to extend Vetter's account to different kinds of counterfactuals, showing some of the challenges accompanying this extension. We identified that one of the main issues concerning Vetter's account of counterfactuals is the definition of the notion of 'being an earlier stage', which was discussed in Chapter VI. Our attempts to fix the nature of the 'earlier stage' relation, in the shape of a 'process of iteration' and two kinds of chain of potentialities, were all met with some problems and to solve them a much longer and deeper discussion would be needed. Therefore, the question about the nature of this relation remained open.

We suggested in Chapter VI that the tentative character of Vetter's account of counterfactuals is at the origin of these issues. The fact that the notion of 'being an earlier stage' is left undefined in Vetter's work means that is open to different interpretations and some of these interpretations could be quite problematic for her theory of possibility.

There are two main risks for Vetter if the 'earlier stage' relation remains undefined and not fully integrated in her metaphysics.

On one side, if this notion is only used in relation to counterfactuals and is left unexplained, then it could seem that its purpose is merely to provide us the semantic results we want concerning counterfactuals. In this way, 'being an earlier stage' would be whatever we need it to be to explain different counterfactuals, with its metaphysical basis modelled entirely for this goal. This would mean that it is the semantics of counterfactuals that guide the development of its metaphysical account, so that this account would risk being *ad hoc*. This would be a metaphysics devised from a semantics, rather than a semantics that looks for its foundations in a metaphysical background. Therefore, if we do not fix what the iteration relation is, Vetter's account of counterfactuals could risk having its metaphysics collapsing into and being biased by its semantics, with this relation turning into a cheap expedient that can change at will depending on our semantic needs.

On the other side, the danger is that if we cannot find other ways to explain the 'earlier stage' relation, then we could conclude that the link at the basis of this relation must be counterfactuals themselves. I have argued that Vetter must explain the link between the potentialities involved in a counterfactual, which is presented in terms of a potentiality 'being an earlier stage' in another. If we fail in our search for such an explanation, then we could end up with a definition of this relation that is itself in terms of counterfactuals, like "A potentiality to be *F* is an earlier stage of an iterated potentiality to be *G* iff *x* could be *G*, if *x* were *F*". This would mean that Vetter cannot reduce counterfactuals to potentialities, and this would be a big blow to her ambition to explain the whole of modality in terms of potentialities. Therefore. if we do not find a viable alternative, and the only way to explain the 'earlier stage' relation existing between potentialities ends up being through counterfactuals, then Vetter cannot reduce every modal notion to potentiality and her project is ultimately jeopardised.

We can see then that the problem of the 'earlier stage' relation cannot be ignored by Vetter. However, I do not think that this problem is unsolvable or that we should inevitably incur the two dangers I presented above. I believe that what is needed is simply more work.

A fuller and complete analysis of iterated potentiality in relation to counterfactuals, which would lead to a real definition of its 'stages', could not only help an account of counterfactuals in terms of potentiality, but could reinforce Vetter's entire picture. Since iterated potentiality is employed as the metaphysical ground of counterfactuals, focusing our attention on these conditionals can only improve our understanding of this kind of potentiality. And given that this kind of potentiality is essential for Vetter's account of possibility as well, the

more complete and clearer picture of it that could come out from this investigation would be extremely beneficial for the whole of her framework. Therefore, a more extensive work on potentiality with a focus on counterfactuals is to be encouraged.

The appeal of an account of counterfactuals fully based on the objects of the actual world and their properties is for me undeniable, given my personal preference for a concrete, actualist and realist view of reality. Given my scepticism towards possible worlds, potentialism looks like a very good alternative which marries well with my metaphysical believes. Therefore, I do wish for Vetter's proposal to work. However, as it stands, the account does need further development, and there are important unanswered questions connected to it. If what left me unsatisfied with the possible world account of counterfactuals was the metaphysics behind it, in the case of a potentiality account of counterfactuals it is the lack of certain details that stops me from fully embracing it, even if I am quite convinced by the general ideas behind potentialism.

Nevertheless, at the same time, this lack of development gives me hope. Since it is not a fully established account yet, there is still the possibility of developing it in new ways and ensuring that the problems are tackled. Vetter's proposal does not need to be final, and it can be the beginning of a profitable and flourishing discussion of counterfactuals in relation to potentiality.

BIBLIOGRAPHY

Adams, Ernest W.,

- 1970, "Subjunctive and Indicative Conditionals", *Foundations of Language*, Vol. 6, No. 1, pp. 89-94.
- 1975, *The Logic of Conditionals*, Reidel Publishing Company, Dordrecht, Holland.

Anderson, Alan Ross, 1951, "A note on subjunctive and counterfactual conditionals", *Analysis*, Vol. 12, pp. 35-38.

Azzano, Lorenzo & Borghini, Andrea, 2020, "Essays on Barbara Vetter's Potentiality – Introduction", *Philosophical Inquiries*, Vol. 8, No. 1, pp. 71-82.

Bennett, Jonathan,

- 1988, "Farewell to the Phlogiston Theory of Conditionals", *Mind*, *New Series*, Vol. 97, No. 388, pp. 509-527.
- 1995, "Classifying Conditionals: The Traditional Way is Right", *Mind, New Series*, Vol. 104, No. 414, pp. 331-354.
- 2003, *A Philosophical Guide to Conditionals*, Oxford University Press, Oxford, England.

Bird, Alexander,

- 2007, *Nature's Metaphysics*, Oxford University Press, Oxford, England.
- 2020, "Possibility and the analysis of dispositions", *Philosophical Inquiries*, Vol. 8, No. 1, pp. 83-96.

Borghini, Andrea, 2016, *A Critical Introduction to the Metaphysics of Modality*, Bloomsbury Academic, London, England.

Brennan, Virginia, 1993, *Root and epistemic modal auxiliary verbs*, Ph.D. thesis, University of Massachusetts.

Carnap, Rudolf, 1947, *Meaning and Necessity*, University of Chicago Press, Chicago, USA.

Casalegno, Paolo S., 1998, *Filosofia del linguaggio: un'introduzione,* Carocci, Roma, Italy.

Casini, Giulia, 2022, "Potentiality and Would-counterfactuals", *Argumenta*, Issue 14. Available at https://www.argumenta.org/article/potentiality-and-would-counterfactuals/.

Chalmers, David J., 2002, "Does conceivability entail possibility?", in Tamar Szabo Gendler & John Hawthorne (eds.), *Conceivability and Possibility*, Oxford University Press, Oxford, England, pp. 145-200.

Chisholm, Roderick M., 1946, "The Contrary-to-Fact Conditional", *Mind*, Vol. 55, No. 220, pp. 289-307.

Contessa, Gabriele, 2016, "Potentiality: From Dispositions to Modality, by Barbara Vetter", *Mind*, Vol. 125, No. 500, pp. 1236-1244.

Crane, Tim, 1996, "Introduction" in D. M. Armstrong, C. B. Martin, & U. T. Place (eds.), *Dispositions. A Debate*, Routledge, London, England, pp. 1-11.

Declerck, Renaat, 2011, "The definition of modality" in A. Patard, & F. Brisard (eds.), *Cognitive approaches to tense, aspect, and epistemic modality*, John Benjamins Publishing Company, Amsterdam, Holland, pp. 21-44.

DeRose, K.

- 1994, "Lewis on 'Might' and 'Would' Counterfactual Conditionals", *Canadian Journal of Philosophy*, Vol. 24, No. 3, pp. 413-18.
- 1999, "Can It Be That It Would Have Been Even Though It Might Not Have Been?", *Philosophical Perspectives*, Vol. 13, *Epistemology*, pp. 385-413.

Divers, John, 2002, *Possible Worlds*, Routledge, London, England.

Dudman, V. H.,

- 1988, "Indicative and Subjunctive", *Analysis*, Vol. 48, No. 3, pp. 113-122.
- 1994, "On Conditionals", *The Journal of Philosophy*, Vol. 91, No. 3, pp. 113-128.

Edgington, Dorothy,

- 1995, "On Conditionals", *Mind,* Vol. 104, No. 414, pp. 235-329.
- 2003, "Counterfactuals and the benefit of hindsight", in Phil Dowe & Paul Noordhof (eds.) *Cause and Chance: Causation in an Indeterministic World*, Routledge, London, England, pp. 12-27.

- 2008, "I – Counterfactuals", *Proceedings of the Aristotelian Society*, Volume 108, Issue 1, part 1, pp. 1–21.

Eagle, A., 2007, "'Might' Counterfactuals", Unpublished, Retrieved from philar-chive.org https://philarchive.org/archive/EAGMCv1 on 02/03/2021.

Ferguson, K.G., 1987, "Truth Conditions for 'Might' Counterfactuals", *The Re-view of Metaphysics*, Vol. 40, No. 3, pp. 483-94.

Fine, Kit, 1975, "Critical notice of *Counterfactuals*", *Mind*, Vol. 84, No. 335, pp. 451-458.

Forbes, Graeme,

- 1985, *The Metaphysics of Modality*, Clarendon Press, Oxford, England.
- 1988, *Languages of Possibility: An Essay in Philosophical Logic*, Blackwell, Oxford, England.

Giannini, Giacomo & Tugby, Matthew, 2020, "Potentiality: actualism minus naturalism equals platonism", *Philosophical Inquiries*, Vol. 8, No. 1, pp. 117-140.

Goodman, Nelson, 1947, "The problem of Counterfactual Conditionals", *The Journal of Philosophy*, Vol. 44, No. 5, pp. 113-128

Harper, W., Stalnaker, R. & Pearce, G. (eds.), 1981, *IFS: Conditionals, Belief, Decision, Chance and Time*, Reidel Publishing Company, Dordrecht, Holland.

Jacobs, J.D. 2010, "A Powers Theory of Modality: Or, How I Learned to Stop Worrying and Reject Possible Worlds", *Philosophical Studies*, Vol. 151, pp. 227-48.

Jubien, Michael, 2007, "Analyzing Modality", *Oxford Studies in Metaphysics*, Vol. 3, pp. 99-139.

Karttunen L. & Peters S., 1978, "Conventional implicature" in Choon-Kyu Oh & David A. Dinneeen (eds.) *Presupposition: Syntax, and Semantics*, Vol. 11, Academic Press, New York, New York, USA, pp. 1-56.

Kratzer, Angelika,

- 1989, "An Investigation of the Lumps of Thought", *Linguistics and Philosophy*, Vol. 12, pp. 607-653.
- 2012, *Modals and Conditionals: New and Revised Perspectives,* Oxford University Press, Oxford, England.

Kripke, Saul,

- 1963, “Semantical analysis of modal logic I. Normal modal propositional calculi”, *Zeitschrift f. math. Logik und Grund. d. Math*., Vol. 9, pp. 67–96.
- 1965, “Semantical analysis of modal logic II. Non-normal modal propositional calculi”, in J. W. Addison, L. Henkin & A. Tarski (eds), *The Theory of Models*, North-Holland Publishing Company, Amsterdam, Holland, pp. 206–220.

Leech, Jessica, 2017, “Potentiality”, *Analysis Reviews*, Vol. 77, No. 2, pp. 457–467.

Lewis, Clarence Irving, Langford, Cooper Harold, 1934, “Symbolic Logic”, *Mind*, Vol. 43, No. 169, pp. 99-109.

Lewis, David,

- 1973a, *Counterfactuals*, Blackwell, Oxford, England.
- 1973b, “Counterfactuals and Comparative Possibility”, *Journal of Philosophical Logic*, Vol. 2, No. 4, pp. 418-446.
- 1979, “Counterfactual Dependence and Time's Arrow”, *Noûs*, Vol. 13, No. 4, *Special Issue on Counterfactuals and Laws* (Nov. 1979), pp. 455-476.
- 1986, *On the Plurality of Worlds*, Blackwell, Oxford, England.
- 1997, “Finkish Dispositions”, *Philosophical Quarterly*, Vol. 47, pp. 143–58.

Martin, C. B., 1994, “Dispositions and Conditionals”, *The Philosophical Quarterly*, Vol. 44, No. 174, pp. 1-8.

McKitrick, Jennifer, 2020, “Potentialities as properties”, *Philosophical Inquiries*, Vol. 8, No. 1, 141-166.

Melia, Joseph, 2003, *Modality*, Acumen Publishing Limited, Chesham, England.

Morato, Vittorio, 2019, “Controfattuali”, *APhEx*, Vol. 20. Available at http://www.aphex.it/index.php?Temi=557D030122027403210506077773 27.

Negri, Sara, 2009, “Kripke completeness revisited”, in Giuseppe Primiero & Shahid Rahman (eds.), *Acts of Knowledge: History, Philosophy and Logic*, *Tributes*, College Publications, London, England, pp. 233-266.

Ramsey, Frank, Plumpton 1931, "General proposition and causality", in R. B. Braithwaite (ed.), *Foundations of Mathematics and other Logical Essays*, Routledge, London, England, pp. 237-255.

Sanford, David H., 1989, *If P, then Q: Conditionals and the Foundations of Reasoning*, Routledge, London, England.

Stalnaker, Robert,

- 1968, "A Theory of Conditionals", in Nicholas Rescher (ed.), *Studies in Logical Theory (American Philosophical Quarterly Monographs 2)*, Blackwell Oxford, England, pp. 98-112.
- 1981, "A defense of conditional excluded middle", in W. Harper, R. Stalnaker, & G. Pearce (eds.), *IFS: Conditionals, Belief, Decision, Chance and Time*, Reidel Publishing Company, Dordrecht, Holland, pp. 87-104.
- 1987, *Inquiry*, A Bradford Book, MIT Press, Boston, USA.

Vetter, Barbara,

- 2011, "Recent Work: Modality without Possible Worlds", *Analysis*, Vol. 71, No. 4, pp. 742-754.
- 2015, *Potentiality: From Dispositions to Modality*, Oxford University Press, Oxford, England.
- 2016, "Counterpossibles (not only) for dispositionalists", *Philosophical Studies*, Vol. 173, pp. 2681–2700.
- 2020, "Replies", *Philosophical Inquiries*, Vol. 8, No. 1, pp. 199-222.

Von Wright, Georg H., 1951, *An Essay in Modal Logic*, North-Holland Publishing Company, Amsterdam, Holland.

Wang, Jennifer,

- 2015, "The Modal Limits of Dispositionalism", *Noûs*, Vol. 49, No.3, pp. 454–469.
- 2020, "Potentiality, modality, and time", *Philosophical Inquiries*, Vol. 8, No. 1, pp. 185-198.

Wildman, Nathan, 2020, "Potential problems? Some issues with Vetter's potentiality account of modality", *Philosophical Inquiries*, Vol. 8, No. 1, pp. 167-184.

Williamson, Timothy, 1998, “Bare Possibilia”, *Erkenntnis*, Vol. 48, pp. 257–73.

Wittgenstein, Ludwig, 1922, *Tractatus Logico-Philosophicus*, International Library of Psychology, Philosophy and Scientific Method, containing the original German, alongside both the Ogden/Ramsey, and Pears/McGuinness English translations, Magdalene College, Cambridge, England.

Wood, Michael, 2003, *Conditionals*, Oxford University Press, Oxford, England.

Yates, David, 2020, “A strange kind of power: Vetter on the formal adequacy of dispositionalism”, *Philosophical Inquiries*, Vol. 8, No. 1, pp. 97-116.