Essays on indiscernibility

Matteo Nizzardo

A thesis submitted for the degree of PhD
at the
University of St Andrews

2024

Identifier to use to cite or link to this thesis:

DOI: https://doi.org/10.17630/sta/844

Candidate's declaration

I, Matteo Nizzardo, do hereby certify that this thesis, submitted for the degree of PhD, which is approximately 75,000 words in length, has been written by me, and that it is the record of work carried out by me, or principally by myself in collaboration with others as acknowledged, and that it has not been submitted in any previous application for any degree. I confirm that any appendices included in my thesis contain only material permitted by the 'Assessment of Postgraduate Research Students' policy.

I was admitted as a research student at the University of St Andrews in September 2019.

I received funding from an organisation or institution and have acknowledged the funder(s) in the full text of my thesis.

Date: December 14, 2023 Signature of candidate:

Supervisor's declaration

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of PhD in the University of St Andrews and that the candidate is qualified to submit this thesis in application for that degree. I confirm that any appendices included in the thesis contain only material permitted by the 'Assessment of Postgraduate Research Students' policy.

Date: December 14, 2023 Signature of supervisor:

Permission for publication

In submitting this thesis to the University of St Andrews we understand that we are giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. We also understand, unless exempt by an award of an embargo as requested below, that the title and the abstract will be published, and that a copy of the work may be made and supplied to any bona fide library or research worker, that this thesis will be electronically accessible for personal or research use and that the library has the right to migrate this thesis into new electronic forms as required to ensure continued access to the thesis.

I, Matteo Nizzardo, confirm that my thesis does not contain any third-party material that requires copyright clearance.

The following is an agreed request by candidate and supervisor regarding the publication of this thesis:

Printed copy

No embargo on print copy.

Electronic copy

No embargo on electronic copy.

Date: December 14, 2023 Signature of candidate:

Date: December 14, 2023 Signature of supervisor:

Underpinning Research Data or Digital Outputs

Candidate's declaration

I, Matteo Nizzardo, understand that by declaring that I have original research data or digital outputs, I should make every effort in meeting the University's and research funders' requirements on the deposit and sharing of research data or research digital outputs.

Date: December 14, 2023 Signature of candidate:

Permission for publication of underpinning research data or digital outputs

We understand that for any original research data or digital outputs which are deposited, we are giving permission for them to be made available for use in accordance with the requirements of the University and research funders, for the time being in force.

We also understand that the title and the description will be published, and that the underpinning research data or digital outputs will be electronically accessible for use in accordance with the license specified at the point of deposit, unless exempt by award of an embargo as requested below.

The following is an agreed request by candidate and supervisor regarding the publication of underpinning research data or digital outputs:

No embargo on underpinning research data or digital outputs.

Date: December 14, 2023 Signature of candidate:

Date: December 14, 2023 Signature of supervisor:

For my sister,
specs of dust in the first light of May.

'I'm going to have a hard time with you two,' said K., comparing their faces yet again. 'How am I to know which of you is which? The only difference between you is your names [...]. So I shall treat you as a single man, and call you both Artur [...]. If I send Artur somewhere you'll both go, if I give Artur a job to do you'll both do it [...]. To me you'll be just one man.'

Kafka, *The Castle*

Abstract

This Thesis is a collection of essays on qualitatively indiscernible entities, i.e. entities which agree with respect to all the qualitative properties they instantiate. In Chapter 1 I introduce various accounts of indiscernibility, and provide a review of the relevant literature. Chapter 2 is dedicated to Leibniz's principle of the Identity of Indiscernibles, the claim that indiscernibility suffices for numerical identity. I argue that if certain assumptions about identity criteria are accepted, the weakest non-trivial interpretation of the principle is one restricted solely to qualitative properties. In Chapter 3 I present a new counterexample to the Identity of Indiscernibles. In Chapter 4 I argue that Anti-Haecceitism, the claim that there are no maximal possibilities which differ only with respect to the non-qualitative possibilities they include, entails that the Identity of Indiscernibles holds of necessity. In Chapter 5 I propose a new account of qualitative properties, according to which a property is qualitative if and only if it is invariant under any identity assignment — where an identity assignment is a function from individuals and worlds to identities. In Chapter 6 I argue that singular reference to indiscernible individuals is possible, and show how current theories of Arbitrary Reference allow for a successful analysis of this phenomenon. In Chapter 7 I defend

Arbitrary Reference against a popular objection, and advance a new probabilistic account of Arbitrary Reference. Finally, in Chapter 8, I show that singular reference to entities to which identity does not apply is impossible.

Acknowledgements

There are many people who, in one way or other, have contributed to this work. First of all, Prof Francesco Berto. The gratitude I owe him extends far more than these few lines could possibly convey. Francesco has been an amazing supervisor and an amazing friend, and has taught me, first and foremost, what it means to be a philosopher. For this I will always be grateful to him. Second, Dr Aaron Cotnoir. In our many meetings, Aaron has set an incredibly high standard of philosophical precision, integrity and humility.

Next, I would like to thank my annual reviewers Prof Crispin Wright, Dr Colin Johnston, and Prof Kevin Scharp, for their comments on the many works which then flowed into this Thesis, as well as the members of the Arché Metaphysics and Logic Research Group. Among them, I especially thank Jace Snodgrass, for all the time he dedicated to discussing philosophy with me, and Stefano Pugnaghi, for reading excerpts of Chapters 6, Chapter 7, and Chapter 8. I also thank Prof Gonzalo Rodriguez-Pereyra for his willingness in discussing with me the main intuitions behind Chapter 2, and Nuno Maia for the many hours he dedicated to challenging my intuitions (and my writing style).

I would like to thank my colleagues in the Philosophy Department of the

University of St Andrews, and the friends I met during my time in Scotland. Again, a special thank you goes to Xintong Wei, Valeria Roberti, Miriam Bowen, Katharina Bernhard, Andrea Oliani, Christopher Masterman, Robert Brown, Francisca Silva, Giulia Schirripa, and Frederik Andersen.

Also, I want to thank my italian friends: Simone Ritarossi, Stefano Rossi, Elisa Frioni, Elisa Chiappini, Luca Delle Cese, Diego Fiorletta, Giacomo Principali, Giorgio Papitto, Massimo Pietrobono, Martina Moriconi, Stefano Pigliacelli, Matteo Scarchilli, Valentino Sperati, Moira Gnesi, Giorgio Mattia, Ludovico Celesti, Fulvio Bernola, Mario Carlo Iusi, Francesco Gerbino, Martina Gerbino, Tiziano Rossi, and Riccardo Piombo. I know most of them since I was young, and I thank all of them for giving me a constant reason to go back home.

Lastly, all this wouldn't have been possible without my family: my father, my mother, and my sister. No words could ever express my gratitude towards them. I want to thank also Dr Claudio Orlando and Prof Tito Magri. In different ways, both of them saved me — and I want them to know that when I find myself thinking about how my life turned out to be, I always, always, think of them.

Funding

This work was supported by the Scottish Graduate School for Arts & Humanities [Grant Number: AHRC: AH/R012717/1]. This work was supported by the St Leonard's College European Doctoral Stipend Scholarship.

Preface

Chapter 4 is published as 'Why I am not an Anti-Haecceitist' (2023). *Synthese* 201(2), 1–14.

Contents

Introduction

Since the first half of the 20th century, qualitative indiscernible entities (indiscernibles for short) have increasingly appeared in distinct and sometimes distant sub-fields of Analytic Philosophy. At first, reasoning about indiscernibles was exclusive of certain enterprises in Metaphysics dealing with the notion of identity, and a possible reduction thereof. Arguing against Leibniz's principle of the Identity of Indiscernibles, A.J. Ayer (1954), Max Black (1952), Peter Strawson (1959) and Robert Adams (1979) have all taken the possibility of indiscernibles seriously, hence undermining one of the most famous reductionist theses about identity. At the same time, the development of Quantum Mechanics introduced indiscernibles in the Philosophy of Physics, prompting animated discussions about the indiscernibility of the universe's smallest components. Entangled electrons, for instance, as well as bosons in the same state of motions, are among the entities which the Received View of Quantum Mechanics sees as eminent examples of actual indiscernibles. (See, among others, French & Krause 2006.) Debates about the extent of particles' indiscernibility are still alive, for many philosophers have opposed the Received View and have suggested distinct strategies to deal with quantum particles and their physical behaviour. Finally, in the

Philosophy of Mathematics, the birth of structuralist accounts of mathematical entities have led to lively debates about mathematical structures containing indiscernibles. (See Shapiro 1997, and Hellman 2004.)

Despite being so widespread across so many philosophical subjects, indiscernibles still raise numerous important questions which meet no agreement. Are indiscernibles metaphysically possible? And if so: are there *actual* indiscernibles? Which properties should we really quantify over in defining a metaphysically substantial notion of indiscernibility? Usually we define indiscernibility through the notion of qualitative properties, but there is no current account of qualitative properties which hasn't been found wanting. Doesn't this jeopardise any attempt to clearly define a relevant notion of indiscernibility? And how can we even know whether, in describing a situation which allegedly contains indiscernibles, we are not just mistakingly talking about a situation with only discernible entities? Can we even really talk about indiscernibles at all?

The Plan

In Chapter 1, I give a general definition of indiscernibility, and review some of the relevant literature in Metaphysics, Philosophy of Physics and Philosophy of Mathematics. I discuss various relations of discernibility, and survey the main reasons why indiscernibles have been (and continue being) approached with circumspection in the philosophical literature.

In Chapter 2 I discuss Leibniz's principle of the Identity of Indiscernibles (PII), according to which if entities x and y are indiscernible, then x and

y are numerically identical. I examine some modern interpretations of this principle, and discuss the so called circularity charge against unrestricted versions of PII. Finally, I present an argument to the extent that if we want PII to afford a reductive analysis of individual-identity, then PII must be restricted to qualitative properties alone.

In Chapter 3 I present a new counterexample to PII using branching worlds: i.e. worlds with multiple incompatible time-lines diverging as a consequence of indeterministic events. I show that my new counterexample is successful against *all* the most common strategies which have historically been employed to defend PII from alleged counterexamples. Following Hawley (2009), these are: the 'identity defense', the 'discerning defense', the 'summing defense' and the 'structure defense'. I conclude that, since all counterexamples to the Identity of Indiscernibles currently on the market are vulnerable to at least one of these strategies, my counterexample puts unprecedented pressure on PII.

In Chapter 4 I establish a connection between PII and Haecceitism, which is the view that there are maximal possibilities which include all the same qualitative possibilities, and yet differ with respect to the non-qualitative possibilities they include. In particular, I argue that if a popular version of the Identity of Indiscernibles relativised to ordinary spatio-temporal entities is not necessarily true, then Haecceitism follows. The main argument of this Chapter, as well as some other arguments in the Thesis, rely on the distinction between 'individuals' (i.e. entities to which identity applies) and 'non-individuals' (i.e. entities to which identity does not apply). Though I do not commit to the existence of non-individuals, I nonetheless make use of the difference between individuals and non-individuals for philosophical purpose.

In Chapter 5 I introduce a new account of qualitative properties, according to which a property P is qualitative if and only if it is invariant under any identity assignment. That is: a property P is qualitative if and only if the fact that P is instantiated by x and not instantiated by y is independent of the fact that the entity z to which x is related in virtue of having P, if any, is indeed z or not. I then develop a formal framework inspired by this account, and test it against different paradigmatic kinds of non-qualitative properties.

In Chapter 6 I challenge the common intuition that singular reference to indiscernible entities is impossible. I suggest that Arbitrary Reference (i.e. the idea that we can refer to individuals with some degree of arbitrariness) allows us to conceptualise and model singular reference to one among many indiscernibles in a clear and consistent manner. I then discuss various theories of Arbitrary Reference and show that they are indeed compatible with the possibility of singular reference to indiscernible individuals.

In Chapter 7, my aim is twofold. First, I argue that the common challenge against Arbitrary Reference according to which it entails that some semantic facts are not grounded in any non-semantic fact is misguided. I argue that friends of Arbitrary Reference can employ indeterministic grounding to show that Arbitrary Reference is compatible with there being no fundamental semantic fact. Second, I advance a new account of Arbitrary Reference as a probabilistic phenomenon, and argue that this new account should be preferred over the more classical versions of Arbitrary Reference developed in Breckenridge & Magidor (2012), Martino (2001), and Woods (2014).

Finally, in Chapter 8, I present four arguments for the impossibility of singular reference to non-individuals: i.e. entities to which identity does not apply.

Chapter 1

Indiscernibility

1.1 Indiscernibles in the Literature

Talk of indiscernible entities is ubiquitous in Philosophy. In Metaphysics and Ontology, indiscernibles are widely mentioned in the contexts of non-reductive accounts of (numerical) identity, of the distinction between qualitative and non-qualitative properties, and of Haecceitism and haecceities.[1] In the Philosophy of Physics, there is an enormous literature about whether particles, elementary and not, are indeed examples of indiscernible entities, both according to Classical Mechanics and Quantum Mechanics. Entangled electrons, as well as bosons in the same state of motion are among the entities that the so-called Received View of Quantum Mechanics acknowledges as eminent examples of actual indiscernibles.[2] In the Philosophy of Mathematics, structuralist accounts of mathematical entities are yet another

[1] See, among others: Black (1952), Cowling (2017), Lowe (2016), and Strawson (1959).

[2] See, among others: Dalla Chiara & Toraldo Di Francia (1993, 1995),Domenech & Holik (2007), and French & Krause (2006)

source of examples involving indiscernibles: the complex numbers, along with all those number structures with non-trivial automorphisms, contain indiscernible places.[3] We find other examples of indiscernibility in debates about non-well founded set theories (e.g. the many distinct Quine atoms allowed in Boffa Set Theory) in sceptical arguments in Epistemology (where some regard the bad and the good cases as phenomenologically and sometimes even evidentially indiscernible), and even in Aesthetics.[4]

In this Section, I will survey the most important examples of indiscernibles in both Metaphysics and Ontology, Philosophy of Physics, and Philosophy of Mathematics.

1.1.1 Metaphysics & Ontology

In Metaphysics and Ontology, indiscernibles have long been at the heart of the debate about the truth of Leibniz's principle of the Identity of Indiscernibles (PII), which claims that qualitative indiscernibility (i.e. equivalence of qualitative properties) is sufficient for numerical identity.

Numerous thought experiments have been proposed against PII, purporting to show that qualitatively indiscernible entities are indeed possible.

Black's Indiscernible Spheres

The most influential counterexample to PII has been suggested by Max Black (1952), who describes a radially symmetrical world containing only two per-

[3]See, among others, Hellman (2004) and Shapiro (1997).

[4]For example of indiscernibles in non-well-founded set theories, see: Aczel (1988), Boffa (1969), and Rieger (2000). For indiscernible cases in Epistemology, see: Rinard (2021), Stroud (1984), and Wright (2004). Finally, for an example of indiscernibles in Aesthetics, see: Danto (1981).

fect iron spheres, one mile in diameter, located at the opposite sides of the world's centre of symmetry, two diameters away from each other. Black imagines the two spheres as having the same physical properties (they have the exact same mass and density, for instance), and the same chemical composition (they are both composed of chemically pure iron). Furthermore, they have the same geometrical properties: they have exactly the same shape, and, among other things, the same volume. Black (1952) argues that there is also no qualitative relational property that can be used to tell them apart: they are both two diameters apart from some perfect iron sphere, and each of them is in the same place as some perfect iron sphere.

The world imagined by Black doesn't need to be infinite, either spatially or temporally. Clearly, we are to imagine the life-span of both spheres as co-extensional with the that of the entire world. Alternatively, we can imagine that the spheres start and stop existing simultaneously. More generally, the supposition is that there is no moment in time when only one of the two spheres exists. It is also perfectly fine for Black's world to be spatially finite, provided it is big enough to contain both spheres, and its centre of symmetry is located exactly where the medians of each of its dimensions intersect.

Although one might think that the space is Black's world cannot be absolute, this is not actually true. Sure, Black only considers spatial relations between objects, but that doesn't mean there is no fact of the matter in Black's universe as where exactly the spheres are located. For the purposes of Black's counterexample, it doesn't make any real difference whether the space is relational, or some location fact (such as: one sphere is in region r_1 and one is in r_2) can be defined. The reason for this is that even if such facts about the exact location of the spheres existed, they wouldn't count as genuinely qualitative facts, for they could discern the spheres only on the basis of the identity of the spatial regions, or points, in question. Furthermore, since Black's world is symmetric, even on the assumption that space is absolute, the spatial region occupied by the first sphere could not even in principle be discerned from the spatial region occupied by the second sphere.

Ayer's Sound Tokens

Another counterexample to PII has been suggested by Ayer (1954, p. 32). Ayer imagines an infinite series of sound tokens:

$$... \; A \; B \; C \; D \; A \; B \; ...$$

with no first or last token, being reproduced at constant intervals of time. Since the description of such situation doesn't seem to involve anything external to the sequence of tokens itself (at least intuitively) then, Ayer claims, we can imagine a world where nothing but these sound tokens are present. In this world, there seem to be no qualitative difference between any two occurrences of the same sound type.

Take, for instance, two successive A sounds. There seems to be no qualitative property that the first token has and the second lacks, or *vice versa*. Both of them must be imagined as to have the same physical properties (i.e. fundamental frequency, amplitude, and Fourier transform) and, being tokens of the same type, they also have the same characteristics in terms of acoustic perception (among which: they have for example the same formants). Plus, they last exactly the same amount of time, so that we cannot distinguish them by their duration. What about their relational properties? Well, they are both followed by a B token, and preceded by a D token. Clearly we can define a notion of temporal distance to try to distinguish them, but it would be of no use in this scenario, for the sequence is infinite — and this means that there is no distance between any given sound token and the *first* sound token in the sequence (or the *last*, for what it matters), for there is no first nor last token. This makes any qualitative relational property instantiated in this scenario symmetric, and therefore non-discerning.

Strawson's Chessboard

A third counterexample to PII is found in Strawson (1959, p. 122), where a world is described consisting of a limited arrangement of black and white squares, resembling a chessboard.

Strawson claims that some of the squares in this world cannot be discerned from one another, for they are symmetrical to each other and lie at the same distance with respect to the squares at the edge of the board. Take for example squares F3 and C6, as depicted in Figure 1.1:

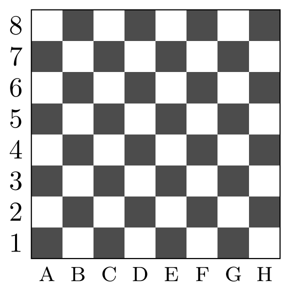


Figure 1.1: Strawson's Chessboard.

They clearly share their non-relational properties, for the only non-relational properties in Strawson's scenario are the color properties 'being black' and 'being white', and clearly bot F3 and C6 are white.[5]

Furthermore, all their qualitative relational properties seem to be the same. For one, they are both 'in the middle' of the chessboard. They both have the property 'having exactly eight neighbouring squares', which distinguishes them both from any square located at the edges of the chessboard (for instance, H1). Also, they both share all the other qualitative relational properties as: 'having some white/black neighbor', 'being one square away from some black/white square' and so on. Clearly, they could be distinguished by means of properties like 'being two squares away from H1', which only F3 has. However, such properties would be non-qualitative, for they would depend on the identity of other squares. Similar qualitative properties, like 'being two squares away from some white square that has only three neighbouring squares', would not distinguish between F3 and C6, for the property 'being a white square that has only three neighbouring squares' is both instantiated by both H1 and A8, and F3 is two squares away from A8, while C6 is two squares away from H1: so both F3 and C6 have the property 'being two squares away from some white square that has only three neighbouring squares'. Finally, notice that in Strawson's world, the number of squares being finite, we will always be able to distinguish, say, F3 from B7. However, we can easily expand Strawson's finite chessboard to an infinite one. Then, all the white squares would be mutually indiscernible, as would all the black

[5]Plausibly, also the property 'being a square' is non-relational, and it is instantiated in Strawson's world. However, both F3 and C6 have it, and therefore it cannot be used to tell them apart. For dissent about the non-relational nature of the property 'being a square', see Allen (2016, pp. 72–77) and Rodriguez-Pereyra (2002, ch. 5).

ones.

1.1.2 Physics & Philosophy of Physics

Indiscernibles appear also in many physical theories. Following French (1989), I discuss two examples offered respectively by Classical Mechanics (CM) and Quantum Mechanics (QM). In both these theories, French argues, particles are thought of as indiscernible. Therefore, *prima facie*, both the ontologies of CM and QM go against PII. (To be precise, they go against certain non-trivial interpretations of PII. More on this in Chapter 2.)

Classical Mechanics

In CM, elementary particles are classified according to some of their physical properties, among which we find mass, spin, charge, and magnetic moment. (This properties are often regarded to be *intrinsic* to the particles, in the sense that they do not depend on the particular state the particle is in.) Particles of the same kind will therefore be indiscernible with respect to these properties. For example, all electrons have a mass of $5.485 \cdot 10^{-4}u$, an electric charge of $-e$, a spin value of $\frac{1}{2}$ and a magnetic moment of $-1.001\mu_B$. On the other hand, protons have a mass of $1.672 \cdot 10^{-27}kg$, an electric charge of e, a spin value of $\frac{1}{2}$ and a magnetic moment of $1.521 \cdot 10^{-3}\mu_B$. Although it is always possible to discern an electron from a proton, whenever we have two or more electrons and we want to tell them apart, the properties with which we classify them will not be enough.

Of course, one can always distinguish elementary particles in CM by means of other properties, like for example their spatio-temporal locations. However, spatio-temporal properties are not only extrinsic, but also non-qualitative. Therefore, elementary particles in CM remain qualitatively indiscernible. Furthermore, the possibility of using spatio-temporal properties to distinguish indiscernible particles rests on the so-called Impenetrability Assumption, according to which no two entities can exist at the same spatial location at the same time. However, French (1989, p. 143) remarks that the Impenetrability Assumption is unwarranted, and that it is still an open question whether it also applies to the quantum domain.

Quantum Mechanics

In QM, the situation is even more complicated. As in CM, particles of the same kind are indiscernible (at least with respect to their intrinsic properties). But according to the so-called Received View of Quantum Mechanics, championed among others by Dalla Chiara and Toraldo di Francia (1993; 1995), Domenech and Holik (2007), and French and Krause (2006), particles in the quantum domain are non-individuals. Following French and Krause (2006, p. 248), we say that a non-individual in the sense of the Received View is an entity to which identity doesn't apply. In particular, we say that if x and y are non-individuals, then sentences like "x is identical x" and "x is distinct from y" are meaningless. (Whether we can use a bit of language like 'x' to refer to a non-individual is an open philosophical question. I will suggest a negative answer to it in Chapter 8.) Proponents of the Received View usually take the non-individuality of elementary particles to follow from the Indistinguishability Postulate, which claims that "[...] there is no way of distinguishing states which differ only by a permutation of the particles." (French, 1989, p. 154) I will discuss this in more detail in Chapter 4.

The Received View is in line with many famous intuitions in the literature about QM. Schrödinger (1953, p. 56), for example, suggests that "[y]ou cannot mark an electron, you cannot paint it red. Indeed, you must not even think of it as marked". Along the same lines, Weyl (1950) remarks that "[e]ven in principle one cannot demand an alibi of an electron". Following this intuition, Dalla Chiara and Toraldo di Francia (1993) write that quantum physics is a "land of anonymity". The Received View of Quantum Mechanics is not universally accepted, though. Many authors, including Bueno (2014), Dorato and Morganti (2013), Berto (2017) and Jantzen (2019), all argue that elementary particles in QM, although intrinsically indiscernible, do still have individuality.

Whether the Received View of Quantum Mechanics is correct doesn't really matter for the purposes of this Thesis. On the other hand, the distinction between individuals and non-individuals, which I will discuss in detail in Chapter 4, will be of the outmost importance in many of the arguments I will present. Finally, it is worth stressing that, although state-independent properties of elementary particles are usually said to be qualitative, a less controversial terminology for the kind of indiscernibility presented in this subsection would be that of *intrinsic* indiscernibility: whereby entities x and y are intrinsically indiscernible whenever they have the same intrinsic properties. In the next subsection we will see yet another kind of indiscernibility: *structural indiscernibility*.

1.1.3 Philosophy of Mathematics

Ante Rem Structuralism is the view that mathematical entities are places in abstract universals-like entities, called *ante rem* structures. (See, among others, Shapiro 1997.) *Ante rem* structures are composed of places and structural properties and relations, and are "[the entities that] the isomorphic models of a theory 'have in common' [...] the form[s] of isomorphic models." (Assadian, 2018, p. 3196)

Within *ante rem* structuralism, places are usually taken to be individuals, and are said to be individuated *only* by (1) their structural properties, and (2) the structural relations they bear to all the other places in a given structure. Since structural properties and relations are 'all there is' to mathematical entities, the kind of indiscernibility we find in *ante rem* structures is restricted to structural properties and relations. We say that places x and y in a given structure are *structurally indiscernible* whenever (1) they have the same structural properties, and (2) they bear the same structural relations to every other places in the structure.[6]

The case of i and $-i$

The above definition of structural indiscernibility is equivalent to the claim that, given a structure S and two places x and y in S, x and y are structurally indiscernible if and only if there is a non-trivial automorphism of S which maps x to y.[7]

[6] For a thorough discussion of structural properties and relations, see Korbmacher & Schiemer (2018).

[7] A function $f : S \to S$ is a non-trivial automorphism of a structure S whenever (1) f is an automorphism of S, and (2) f is not the identity function.

Consider the field of complex numbers $\mathbb{C}$, and the function $f : \mathbb{C} \to \mathbb{C}$ which maps any complex number $x + iy$ to the complex number $x - iy$. As Ladyman (2005, p. 219) points out, f is a non-trivial automorphism of the complex field, and since the structure of the field of complex numbers is preserved under f, it follows that i and $-i$ must have the same structural properties and must stand in the same structural relations to all the other places in the complex field. The case of i and $-i$ in the complex numbers structure is a case of structural indiscernibility.

The Cardinal Two Structure

Another instance of structural indiscernibility is represented by the places in the Cardinal Two Structure, depicted in Figure 1.2:

Figure 1.2: Cardinal Two Structure.

This structure is composed of two places, x and y, with no structural properties and no structural relations. Therefore, it is easy to see that there is a structure preserving function f which maps x to y. The same happens with other cardinal structures, as long as they contain more than one place. That is: since x and y are structurally indiscernible in virtue of there being no structural property or relation in the Cardinal Two Structure, it follows that any two places in any higher cardinal structures will exhibit the same kind of indiscernibility.

This is the reason why Hellman (2004, p. 572) remarks that the finite cardinal structures of the *ante rem* structuralist are the "[...] ultimate offence against Leibnizian scruples". I will discuss Hellman's remarks about cardinal structures and a structuralist version of the Identity of Indiscernibles in Chapter 6.

1.2 Degrees of Discernibility

Since Quine (1976), it is customary to distinguish between three degrees of discernibility: *absolute discernibility*, *relative discernibility*, and *weak discernibility*.[8]

1.2.1 Absolute Discernibility

Entities x and y are absolutely discernible whenever there is a property P which only one between x and y instantiates. Consider again Donald Trump and Joe Biden, who are respectively 6.3 and 6.0 feet tall. They are absolutely discernible, for there is at least one property, say 'being 6.3 feet tall', which only one of them (i.e. Trump) instantiates.

Before defining absolute discernibility (or any other degree of discernibility) in a formal way, we need to define the notion of 'discernibility in a structure' (Ladyman et al., 2012, p. 166–167). Say that a *structure* is composed of four elements: (1) a collection of individual entities, which we call *domain*, (2) a collection of distinguished elements of the domain — we call these elements *constant elements*, (3) a collection of n-ary relations on the

[8]This terminology is quite recent. Quine (1976, p. 113–114) originally called these degrees of discernibility (or 'grades of discriminability'): *strong discriminability*, *moderate discriminability*, and *weak discriminability* respectively. More recent literature, however, has settled for a different terminology. (See, among others, Caulton & Butterfield 2012 and Ladyman et al. 2012.)

domain for every n in $\mathbb{N}$, and (4) a collection of n-ary functions on the domain for every n in $\mathbb{N}$.[9]

Say also that each structure A has a signature A_s, i.e. a set of symbols such that: (1) for any constant element in A there is a constant symbol in A_s, (2) for any relation in A there is a relation symbol in A_s, and (3) for any function in A there is a function symbol in A_s. Ladyman, Linnebo and Pettigrew (2012, p. 166) remark that a structure A together with its associated signature A_s uniquely determine four distinct first-order languages: two with the identity symbols and two without it.[10] Call $\mathcal{L}_A$ the first-order language without identity determined from A and A_s, such that any formula φ in $\mathcal{L}_A$ is such that any constant symbol, relation symbol and function symbol appearing in φ is already in A_s.

Relative to $\mathcal{L}_A$, Ladyman, Linnebo and Pettigrew (2012, p. 167) define absolute discernibility *in a structure* as follows. Let A be a structure, and a and b elements of A's domain. Then, we say that a and b are absolutely discernible in A relative to $\mathcal{L}_A$ if and only if there is a formula $\varphi(x)$ in $\mathcal{L}_A$ (that is: a formula with variable x free) such that $\varphi(a)$ is modeled by A and it is not the case that $\varphi(b)$ is modeled by A. (Here and in the following, '$\varphi(a)$ is modeled by A' abbreviates: A models $\varphi(x)$ under an assignment mapping 'x' to a.) Clearly, we can give other definitions of absolute discernibility in A if we consider other languages, be them among the ones determined from A and A_s or not. Since this is an introduction, however, I will not go into

[9] This and the following definitions are borrowed by (Ladyman et al., 2012, p. 166–167).

[10] For the present purposes, we do not have to go into much details about these four languages. The only thing we need here is a reference language (preferably among these four) to which relativise our definitions of discernibility in a structure. The interested reader is referred to Ladyman et al. (2012, p. 166–171).

further details.

1.2.2 Relative Discernibility

Entities x and y are relatively discernible whenever there is a two-place relation R such that x stands in R to y but y doesn't stand in R to x. An example of relatively discernible entities is suggested in Quine (1976, p. 113): any two ordinal numbers x and y are relatively discernible given the relation 'being less than —'. This is because since x and y are ordinal numbers, then either it is the case that x is less than y or it is the case that y is less than x. However, the relation 'being less than —' is asymmetric: if it is the case that x is less than y, than it cannot be the case that y is less than x.

Keeping in mind the technicalities involved in the formal definition of absolute discernibility in a structure, and following again Ladyman et al. (2012, p. 167), we can define relative discernibility in a structure as follows. Let A be a structure, and a and b elements of A's domain. We say that a and b are relatively discernible in A relative to $\mathcal{L}_A$ if and only if there is a formula $\varphi(x, y)$ in $\mathcal{L}_A$ (that is: a formula with variables x and y free) such that $\varphi(a, b)$ is modeled by A and it is not the case that $\varphi(b, a)$ is modeled by A. (Recall: $\mathcal{L}_A$ is the first-order language without identity determined from A and A_s, such that any formula φ in $\mathcal{L}_A$ is such that any constant symbol, relation symbol and function symbol appearing in φ is already in A_s.)

1.2.3 Weak Discernibility

Entities x and y are weakly discernible whenever there is a two-place relation R such that x stands in R to y but x doesn't stand in R to itself. Examples of weakly indiscernible entities are, according to Ladyman, Linnebo and Pettigrew (2012, p. 165): entangled fermions, which can be weakly discerned by the relation 'having opposite spin to —', and the complex numbers i and $-i$, which can be weakly discerned by the relation 'being the additive inverse of —'.[11]

Ladyman, Linnebo and Pettigrew (2012, p. 167) define weak discernibility in a structure as follows. Let A be a structure, and a and b elements of A's domain. Then, a and b are weakly discernible in A relative to $\mathcal{L}_A$ if and only if there is a formula $\varphi(x, y)$ in $\mathcal{L}_A$ (that is: a formula with variables x and y free) such that $\varphi(a, b)$ is modeled by A and it is not the case that $\varphi(a, a)$ is modeled by A.[12]

One important result found in Ladyman et al. (2012) about weak discernibility is that weak discernibility is the most discerning natural non-trivial discernibility relation. The sense in which weak discernibility is the *most discerning* among the non-trivial discernibility relations is the following. Given two discernibility relations R_1 and R_2 we say that R_1 is *less discerning* than R_2 (given a structure A and a language $\mathcal{L}$) if and only if: (1) any two

[11] Saunders (2006, p. 58–60) notes that although weak discernibility works with entangled fermions (which always have spin opposite to each other in accordance with Pauli's Exclusion Principle), it doesn't always work with entangled bosons, for some pairs of entangled bosons have exactly the same spin.

[12] Recall, again, that $\mathcal{L}_A$ is the first-order language without identity determined from A and A_s, such that any formula φ in $\mathcal{L}_A$ is such that any constant symbol, relation symbol and function symbol appearing in φ is already in A_s.

individuals a and b in A's domain which are R_1-discernible in A relative to $\mathcal{L}$ are also R_2-discernible; and (2) there are at least two individuals a and b in A's domain, which are R_2-discernible in A relative to $\mathcal{L}$ but not R_1 discernible.

As we will discuss in detail in Chapter 3, weak discernibility has played a major role in recent debates about PII. In particular, some authors have tried to defend PII against alleged counterexamples on the basis of finding relations which would weakly discern the supposedly indiscernible individuals at hand. One example is Black's radially symmetrical universe (already discussed in Section 1.1.1). Among others, Caulton and Butterfield (2012, p. 50) have proposed that Black's spheres are far from indiscernible. In fact, they can be weakly discerned via the irreflexive relation 'being two diameters apart from —', in which each sphere stands to the other sphere, and yet not to itself.

Another example comes from the case of the complex numbers i and $-i$ in the context of *ante rem* structuralism. (I have already presented this case in Section 1.1.3.) The problem that i and $-i$ raise for *ante rem* structuralism can be quickly put as follows. If *ante rem* structuralism is correct, then there is nothing more to numbers than their structural properties and relations. However, this yields a version of PII whereby no two numbers can be structurally indiscernible. Therefore, the *ante rem* structuralist seems forced to identify i with $-i$. (I will discuss this in more details in Chapter 6.) Ladyman (2005) proposes to rescue *ante rem* structuralism by holding that the relation 'being the additive inverse of —' can be used to weakly discern i and $-i$, thereby avoiding the identification of the two numbers on pain of the violation of PII.

Strategies like these, however, are not unproblematic. Caulton and Butterfield's (2012) defense of PII against Black's counterexample is challenged among others by Lowe (2016, p. 53–58), who argues that the alleged weak discernibility of Black's spheres claimed by Caulton and Butterfield is ultimately just a byproduct of the formal notation they use. On the other hand, Ladyman's (2005) defense of *ante rem* structuralism is challenged by MacBride (2006, p. 67), who remarks that its success depends on whether "[...] the obtaining of irreflexive relations presupposes, in some relevant ontological sense, the numerical diversity of the objects they relate, or not". A similar worry is echoed in French (2019), who argues that appealing to irreflexive relations to ground the individuality of the entities that stand in those very relations might be circular: for "[...] in order to appeal to such relations, one has had to already individuate the [entities] which are so related and the numerical diversity of the [relevant entities] has been presupposed by the relation which hence cannot account for it".

1.2.4 Intrinsic Discernibility

To absolute, relative, and weak discernibility, Ladyman, Linnebo and Pettigrew (2012) add what they call *intrinsic discernibility*, which they define as follows: entities x and y are intrinsically discernible whenever there is some intrinsic property P such that only one between x and y instantiates P.

Formally: let A be a structure, and a and b elements of A's domain. Then, a and b are intrinsically discernible in A relative to $\mathcal{L}_A$ if and only if there is a formula $\varphi(x)$ in $\mathcal{L}_A$ without quantifiers and constants such that

$\varphi(a)$ is modeled by A and it is not the case that $\varphi(b)$ is modeled by A. (See Ladyman et al. 2012, p. 167.)

Following Caulton and Butterfield (2012), Ladyman, Linnebo and Pettigrew (2012, p. 167) say that a property P is intrinsic to some individual entity x "[...] if the existence and nature of other [individual entities] is counterfactually irrelevant to [x] having [P]". This is the reason why, in their definition of intrinsic discernibility in a structure, the discerning formula $\varphi(x)$ is required to be devoid of both quantifiers and constants.[13].

It is easy to see that intrinsic discernibility is the least discerning among the discernibility relations defined so far. Ladyman, Linnebo and Pettigrew (2012, p. 167) also show that intrinsic, absolute, relative and weak discernibility are increasingly discerning. This means that, given a structure A and a language $\mathcal{L}$, for any two entities a and b in A: (1) if a and b are intrinsically discernible in A relative to $\mathcal{L}$, then they are also absolute discernible; (2) if a and b are absolutely discernible in A relative to $\mathcal{L}$, then they are also relatively discernible; and (3) if a and b are relatively discernible in A relative to $\mathcal{L}$, then they are also weakly discernible. (See also Ketland 2011.)

1.2.5 Other Kinds of Discernibility

Similar definitions are given by Caulton and Butterfield (2012). Unlike Ladyman, Linnebo, and Pettigrew (2012), Caulton and Butterfield restrict their attentions to languages without individual constants and function symbols (i.e. languages whose non-logical vocabulary consists only of predicates and

[13]For a detailed discussion of intrinsic and extrinsic properties, see: Cameron (2009), Eddon (2010), Hoffmann-Kolss (2010), and Lewis (1983)

relation symbols).

Caulton and Butterfield (2012) characterise *absolute discernibility* as a disjunction of two distinct kinds of discernibility. For them, entities a and b are absolutely discernible in a given structure A relative to a language $\mathcal{L}$ (without constants and function symbols) whenever either (1) a and b are intrinsically discernible in A relative to $\mathcal{L}$, or (2) a and b are externally discernible in A relative to L.[14]

Given a structure A and some entities a and b in A, we say that a and b are *intrinsically discernible* in A relative to a constantless and functionless language $\mathcal{L}$ whenever there is a one-place $\mathcal{L}$-formula $\varphi(x)$ with no bound variables which only applies to one of a and b. As Caulton and Butterfield (2012, p. 47) explain, the definition of intrinsic discernibility applies both to primitive one-place predicates like $F(x)$, and to formulas the are obtained by replacing all the places of a n-places relation symbol with all occurrences of the same variable, like in the case of $H(x, x, x)$ — provided of course the resulting formula doesn't contain any bound variables.[15]

Given a structure A and some entities a and b in A, we say that a and b are *externally discernible* in A relative to a constantless and functionless language $\mathcal{L}$ whenever there is a one-place $\mathcal{L}$-formula $\varphi(x)$ with bound variables that only applies to one of a and b. Unlike intrinsic discernibility, external discernibility "[...] follows from the relations the two [individuals] a and b

[14]Since $\mathcal{L}$ doesn't have any individual constants, in what follows 'a' and 'b' must be considered as names in the meta-language.

[15]This, Caulton and Butterfield explain, is supposed to capture the intuitive idea that intrinsic discernibility obtains in virtue of some property or relation whose instantiation only depends on how the instantiating individuals are in themselves, i.e. it doesn't depend on how other individuals are.

have to *other* [individuals]." (Caulton & Butterfield, 2012, p. 48.)

Caulton and Butterfield then characterise *relative discernibility* and *weak discernibility* as follows. Given a structure A and some entities a and b in A, we say that a and b are *relatively discernible* in A relative to a constantless and functionless language $\mathcal{L}$ whenever there is a $\mathcal{L}$-formula $\varphi(x, y)$ with two free variables which is satisfied by a and b only in one order. We say instead that a and b are *weakly discernible* in A relative to a constantless and functionless language $\mathcal{L}$ whenever there is a $\mathcal{L}$-formula $\varphi(x, y)$ with two free variables which is satisfied by a and b in any order, but not by a, or b, taken twice. (Notice that Caulton and Butterfield's definitions of relative and weak discernibility are almost the same as Ladyman, Linnebo, and Pettigrew's definitions.)

Let $\mathcal{L}$ be a first-order language without individual constants, and let D be a domain of quantification such that $\mathcal{L}$'s predicates are interpreted as subsets of D^n (according to the degree of the predicates). Following Caulton and Butterfield (2012, p. 42–43), call D together with the assignment of an extension to each of $\mathcal{L}$'s predicates a *structure*. Let π be a permutation on D: i.e. a bijection from D to itself. Then we say that π is a symmetry if and only if, for any predicate P^n of $\mathcal{L}$, the extension of P^n is invariant under π.

Caulton and Butterfield (2012, p. 50–62) prove some important results about absolute indiscernibility, defined as discernibility by means of either intrinsic or external formulas. First, they show that, for any structure $\langle D, i\rangle$, if a permutation π on D is a symmetry, then equivalence classes of absolute indiscernibility are invariant under π. (In Caulton and Butterfield's terminology, two individuals a and b are *absolutely indiscernible* whenever they are *not* absolutely discernible.) It follows from this that: (1) if there is an individual a in the D which is absolutely discernible from all the other individuals in D, then $a = \pi(a)$ for every symmetry on D; and that (2) if all individuals in D are pairwise absolutely discernible, then the only symmetry on D is the identity map.[16] The converse of Caulton and Butterfield's first result doesn't hold: there are structures with permutations which are not symmetries and yet preserve the absolute indiscernibility classes. Finally, it is possible to show that if D is finite and a and b in D are absolutely discernible, then there is a symmetry π such that $\pi(a) = b$.[17]

1.3 Scepticism about Indiscernibles

Indiscernibles are not easy to deal with: concede that they are there, and you will immediately be flooded with philosophical issues that are painstakingly hard to solve. According to Assadian (2019), there are three main kinds of scepticism about utter indiscernibles (i.e. entities that are not even weakly discernible): an ontological kind of scepticism, an epistemic kind of scepticism, and a linguistic kind of scepticism.[18] According to Assadian (2019, p.

[16] The readers who are familiar with Caulton and Butterfield's (2012) work will notice that, in surveying their results, I adopt a slightly different terminology. This is mainly due to the fact that, following Saunders and Muller (2008), Caulton and Butterfield (2012, p. 31) define individuality in terms of absolute discernibility: where an 'object' is the potential referent of a name, an 'individual', in Caulton and Butterfield's sense, is an object that is absolutely discernible from any other object. On the contrary, following French and Krause (2006), I say that an 'individual' is any entity to which identity applies. Unlike Caulton and Butterfield's definition, therefore, my definition of individuality doesn't exclude the existence of individuals which are absolutely indiscernible.

[17] Another source of interesting results about different degrees of discernibility is Button (2017).

[18] Assadian's (2019) main aim is to defend utter indiscernibles from these kinds of scepticism, by showing that all the arguments against utter indiscernibles easily generalise to

2553) examples of utter indiscernibles include the places in the Cardinal Two Structure, discussed in Section 1.1.3, as well as pairs of unentangled bosons in direct product states.

In this Section I discuss the three kinds of scepticism about indiscernibles outlined in Assadian (2019), to wit: ontological scepticism, epistemic scepticism, and linguistic scepticism. To these I add another kind of scepticism, which I call 'meta-theoretical scepticism'. With this, I aim to give the reader a sense of the many philosophical issues indiscernibles give raise to, as well as a sense of why many authors still remain unconvinced of their existence. (See, among others: Della Rocca (2005), Hellman (2004), and Rodriguez-Pereyra (2006).) In the following Chapters, I will directly engage with (at least) two of these kinds of scepticism. Against ontological scepticism, I will argue in Chapter 3 that branching worlds afford the resources to construct very strong counterexamples to the principle of the Identity of Indiscernibles, i.e. the thesis that for any individuals x and y, if x and y share all the same qualitative properties, then x and y are identical. (I discuss the Identity of Indiscernibles in detail in Chapter 2.) Finally, against linguistic scepticism, I will argue in Chapters 6 and 7 that we can obtain singular reference to indiscernible entities via Arbitrary Reference, i.e. the thesis that in certain circumstances, we can refer to individuals with some degree of arbitrariness.

arguments against less philosophically controversial entities, like weakly discernible entities and relatively discernible entities. By showing that there is no single philosophical issue that only utter indiscernibles give rise to, Assadian aims to rehabilitate indiscernibles as genuine entities. A similar argument, to the extent that if weakly discernible entities exist, then utter indiscernible entities exist too, can be found in Hawley (2006).

1.3.1 Ontological Scepticism

The first kind of scepticism about utter indiscernibles can be seen as a form of scepticism with respect to their existence (hence the name 'ontological scepticism'). Why should we doubt, or be wary of, the existence of such entities? For one, notice that utter indiscernibles are incompatible with any reductive general account of identity. If x and y are utterly indiscernible, then x and y are numerically distinct and there is no property or relation, however complex, that can be used to discern them (at least, non trivially). Any property of x is in fact also a property of y and *vice versa*, and any relation (however complex) in which x stands to any other entities is also a relation in which y stands to the same entities, and *vice versa*. Therefore, the only possible answer to the question of what grounds, or explains, the numerical distinctness of x and y, is that it is a fundamental fact, not further explainable, that x and y are indeed not the same object. That is, there are no other facts of the matter one can appeal to even in principle that can explain the fact that x and y are indeed distinct: this is where explanations end.

It follows that if utter indiscernibles exist, then identity is fundamental.[19] Commitment to fundamental identity and non-identity facts, however, is, as Della Rocca (2005) points out, quite unpalatable. For one, if identity is

[19]This, at least, at a general level. For suppose we know that the only utterly indiscernible entities there are are indeed places in mathematical structures and entangled bosons. In this case, one might want to say that although it is true that the identity of certain abstract entities and certain elementary particles is indeed fundamental, this doesn't entail that identity is fundamental *across the board*. We might still be able to correctly ground, or explain, the identity of non fundamental physical objects, for instance, by reference to the properties and relations which discern them.

fundamental, then there is no principled reason why we should exclude, for example, that instead of having just one left hand, we have twenty, or two hundred, utterly indiscernible co-located left hands.

Explaining (or grounding) identity is, according to Della Rocca, one of the main advantages of accepting some non trivial version of the Identity of Indiscernibles. (More on this in Chapter 2.) For if we hold some non trivial version of PII then we can ground the identity of x and y in their indiscernibility with respect to the properties we hold relevant for their distinctness. In other words, we can replace any instances of the identity-involving formulas '$x = y$' and '$x \neq y$' with instances of the identity-free formulas '$\forall_{\mathcal{K}} P(Px \leftrightarrow Py)$' and '$\exists_{\mathcal{K}} P((Px \wedge \neg Py) \vee (Py \wedge \neg Px))$', respectively.[20]

As Assadian (2019, p. 2553) correctly remarks, PII is the cornerstone of virtually every reductive account of identity, for any account according to which the identity of individuals is not fundamental will explicitly endorse, or implicitly entail, some non trivial version of PII. One notable example is the so-called 'Bundle Theory of Substance', which comes in two strands. According to the first, individuals are bundles of co-instantiated universals.[21] (See Hawthorne 1995.) According to the second, individuals are bundles of co-instantiated tropes. (See Williams 1966.) It should be easy to see how these two strands of the Bundle Theory of Substance, which ground the identity of individuals in their universals and tropes respectively, entail two quite

[20] I use the symbols '$\forall_{\mathcal{K}}$' and '$\exists_{\mathcal{K}}$' to express the fact that the relevant quantification is restricted to a certain kind $\mathcal{K}$ of properties. I say that the two formulas with which we can substitute '$x = y$' and '$x \neq y$' are identity-free for otherwise the version of PII in question would be trivial. (I will discuss this in details in Chapter 2.)

[21] Here the term 'universal' is intended to exclude haecceitistic properties, like the property 'being Napoleon'. For more on universals, see: Armstrong (1978). For more on haecceitistic properties, see: Adams (1979), and Hawthorne (2003).

different versions of PII. In particular, the first strand, according to which individuals are bundles of co-instantiated universals, entails that if individuals x and y have the same universals, then they are the same individual. On the other hand, the second strand, according to which individuals are bundles of co-instantiated tropes, entails that if individuals x and y have the same tropes, then they are the same individual.

Aside from the issue of grounding identity facts in facts that do not involve identity, Assadian (2019, p. 2557) notices that some authors seem to be ontologically sceptical about utter indiscernibles due to the belief that for an entity x to be said to be an 'object', it is necessary that x has some properties or stands in some relations that can discern it from any other objects.[22] However, Assadian argues that this form of scepticism about utter indiscernibles is unjustified under any conception of 'object', for all principled definitions of objecthood will inevitably entail that if there is a problem with utter indiscernibles, then there is a problem with weakly discernible entities and sometimes even relatively discernible ones.

1.3.2 Epistemic Scepticism

The second kind of scepticism about utter indiscernibles is epistemic in nature. This kind of scepticism questions how we can come to know that utterly indiscernible entities are indeed numerically distinct. The idea is that we usually explain our knowledge of the numerical distinctness of two entities by appealing to certain properties and relations with which we can discern them. For example: how do we explain that we know that Donald Trump is numeri-

[22]Examples are: Button (2006), Hellman (2007), and Saunders (2003).

cally distinct from Hillary Clinton? One natural answer is that we know that Donald Trump is numerically distinct from Hillary Clinton *because* we know that Donald Trump, and not Hillary Clinton, was President of the United States. ('Having been the President of the United States' is a property that we can use to discern Donald Trump from Hillary Clinton.) Informally, this is an instance of the very intuitive principle according to which if you want to know whether two individual are numerically distinct, all you have to do is to look for any feature that only one of them has: if you find it, then you know that they are indeed distinct. Formally, it is an application of Leibniz's principle of the Indiscernibility of Identicals, which states that for all individuals x and y, if x is identical to y, then all properties of x are also properties of y (and *vice versa*), and all relations x stands in to any other individuals are also relations y stands in to exactly the same individuals (and *vice versa*).[23]

If finding a discerning feature is the only way we can come to know whether entities a and b are numerically distinct, then we can never 'come to know' whether two utter indiscernibles are indeed distinct or not. (Or we can never explain how we know that two utter indiscernibles are distinct.) And this is because if a and b are utterly indiscernible, then there is no feature which can discern them. But then: if utter indiscernibles exist, how can we "[...] explain our knowledge of their numerical diversity?" (Assadian 2019, p.

[23]The Indiscernibility of Identicals is virtually universally accepted as a principle governing numerical identity, and it is oftentimes used in the metaphysical literature with far-reaching and counterintuitive results. (See among others: Evans (1978) on indeterminate identity, Heil (2003, p. 9–10) on the reducibility of propositions to sets, and Kripke (1971) on the necessity of identity.) For an in-depth analysis of these controversial arguments from the Indiscernibility of Identicals, see Magidor (2011).

2554.) Assadian points out that, pretty much like the ontological scepticism discussed in Section 1.3.1, epistemic scepticism about utter indiscernibles exports easily to weakly discernible entities, at least when the entities we are talking about are abstract. In particular, Assadian (2019, p. 2559) argues that, in the case of abstract entities:

> [...] whichever explanation we choose to account for the reliability of our beliefs about abstract objects which are, somehow, *discernible* would be equally adequate to account for the reliability of our beliefs about utter indiscernibles.

A case in point are the complex numbers i and $-i$, discussed in Section 1.1.3. Assadian argues that the reliability of our beliefs about the numerical distinctness of i and $-i$ is a function of the reliability of our beliefs in the axioms that govern complex numbers. We believe in the axioms of complex numbers, and our beliefs about them are reliable. From those axioms, it follow deductively that there are exactly *two* square roots of -1: i and $-i$. Therefore, our true belief that i and $-i$ are indeed numerically distinct is reliable, and this reliability doesn't involve having found a feature that only one between i and $-i$ has. The same can be said, according to Assadian, with respect to other structures with utter indiscernibles, like the Cardinal Two Structure discussed in Section 1.1.3.

1.3.3 Linguistic Scepticism

A third kind of scepticism about utter indiscernibles is linguistic. Suppose that utter indiscernibles exist: the linguistic sceptic now challenges us to explain how we can singularly refer to them, as we seem to do when we say: "if a and b are utterly indiscernible, then a has all the same properties as b". A natural thought is that since a and b are indiscernible, then when we use the term 'a' to refer to a, there is nothing in our use of the term that indeed determines that it refers to a as opposed to b. We cannot impose further constraints on our referring expressions to 'make sure', as it were, that our term 'a' refers to a. And this is because no constraint will never be able to single out a and not b.

Compare this with the case of the 'almost indiscernible spheres', as found in Adams (1979, p. 17). If you have two otherwise indiscernible spheres except for the fact that one has a little chemical impurity that the other lacks, you can impose a constraint on your referring expressions. For example, you might say: "with 'a', I intend to refer to the sphere with the chemical impurity", and that would do the trick. But in the case of utterly indiscernible entities, like the places in the structuralist's cardinal structures, there seem to be nothing that determines the reference of any singular term standing for any one of such entities. About the places in the structuralist's cardinal structures, Hellman (2004, p. 572) asks:

> How is it that any [place] is distinct from any other? Indeed, how can we make sense of referring to any one of them as opposed to any other, or mapping any one of them to or from anything else [...]?

Questions of this kind, however, arise also for entities that are not utterly indiscernible. Consider again Black's spheres, discussed in Section 1.1.1.

Despite being qualitatively indiscernible, Black's spheres are not utterly indiscernible: as we saw in Section 1.2.3, they are weakly discernible given the irreflexive relation 'being two diameters apart from —', in which each sphere stands to the other sphere, and yet not to itself. Despite being weakly discernible, Black (1952) is sceptical about the possibility of referring to them singularly. In a famous passage (Black 1952, p. 156), he argues:

> How can I [consider only one of my spheres and designate it as 'a'], since there is no way of telling them apart? *Which* one do you want me to consider? [...] I don't know how to identify one of two spheres supposed to be alone in space and so symmetrically placed with respect to each other that neither has any quality or character the other does not also have.

So again, the sceptical challenge fails to address utter indiscernibles alone. Interestingly, Assadian agrees with Hellman about the impossibility of singular reference to utter indiscernibles, and suggests that the terms we use when we try to singularly refer to indiscernible entities are not genuine singular terms. (Assadian 2019, p. 2560.) I will challenge this intuition in Chapter 6 of my Thesis, where I will argue that genuine singular reference to indiscernible is possible.

1.3.4 Meta-theoretical Scepticism

To Assadian's three kinds of scepticism I want to add a fourth one, which I call meta-theoretical scepticism. This kind of scepticism about utter indiscernibles arises from the fact that despite our theories might involve quantification over indiscernibles, in our usual model theoretic practices our mathematical models will only contain discernible entities. This means that quantifications over indiscernibles in our theories will be usually interpreted as quantifications over discernible entities in the models.

This happens because we usually define our models as set-theoretic entities, which are either ZFC sets or ZFU sets. This means that either our models only contain pure sets, or they contain sets and urelements alike (i.e. non set-like entities which can be members of sets). In the first case, the interpretations of the individual constants and the assignments of the variables in the formulas of our (relevant) theory will be sets in the ambient theory of ZFC. And since in ZFC any two distinct sets are discernible, then any two constants standing for distinct indiscernibles will always be assigned discernible elements of the domain (provided they are indeed assigned *distinct* elements of the domain). That any two distinct sets in ZFC are discernible follows from the fact that, in ZFC, if x and y are numerically distinct, then there is a set z such that only one between x and y is a member of z.[24]

In case our models contain urelements, even if it is true that we can consider any two urelements as indiscernible in the sense that any permutation of urelements can be extended to an automorphism of the entire domain, it is also true that also urelements (and not only sets) are subject to the way the underlying logic characterises identity. This means, Krause and Coelho

[24] This is a consequence of the following theorem of ZFC: $\forall x \forall y(\forall z(x \in z \leftrightarrow y \in z) \rightarrow (x = y))$. To prove this statement is an easy task. Suppose x and y are sets. Suppose further that for all sets z, $x \in z$ if and only if $y \in z$. By Universal Instantiation, $x \in \{x\}$ if and only if $y \in \{x\}$. Since $x \in \{x\}$ by definition, then also $y \in \{x\}$. But that just means that $x = y$. This statement figures also in Fraenkel, Bar-Hillel and Lévy's (1973, p. 28) definition of set identity, whereby $(x = y)$ is defined by the formula: $\forall z(z \in x \leftrightarrow z \in y) \wedge \forall z(x \in z \leftrightarrow y \in z)$.

(2005, p. 197) argue, that urelements cannot really be taken as genuinely indiscernible. For independently of whether our ambient theory is a first order or second order set theory, it will still be the case that for any urelements x and y, if x and y are numerically distinct, there will be a set which separates them, i.e. a set which contains only one between x and y.

One way to understand the reasons behind this model-theoretic form of scepticism is the following: it seems that as long as we are talking about individuals (i.e. entities which have determinate identity conditions), we can 'correctly' model indiscernibles by means of ordinary discernible entities (which we embed in structures over which we define operations or conditions to make their elements *look* indiscernible), and therefore we don't really need indiscernibles in our model building practices after all.

Chapter 2

The Identity of Indiscernibles

2.1 Introduction

In this Chapter I discuss the principle of the Identity of Indiscernibles (PII), one of the most controversial principles connecting the notions of identity and indiscernibility. After a brief history of the principle (Section 2.2.1), I discuss some of the modern interpretations of PII (Section 2.2.2), focusing in particular on Adams' and Wiggins' understandings of the principle (Section 2.2.3). I compare PII with other important identity criteria, among which are the Axiom of Extensionality in ZF(C) and Davidson's (1969) criterion of event-identity, and discuss the so-called circularity charge against unrestricted versions of PII (Section 2.3). Here I argue for two claims. The first one is that it is possible to distinguish between two distinct kinds of circularity displayed by unrestricted versions of PII: a strong circularity and a weak one (Section 2.3.1). The second one is that, if one believes PII should not be used to explain, or define, individual-identity, then there is no reason why they should not endorse a weakly circular PII (Section 2.3.2). Finally, I argue that if instead one believes that PII should be understood as a reductive thesis about individual-identity, then no version of PII quantifying over non-qualitative properties is a viable option (Section 2.4). I do this by discussing Rodriguez-Pereyra's (2006) version of PII according to which no two entities can instantiate the same non-trivializing properties, and showing that any scenario in which only this version of PII is true either displays an infinite regress in the explanation of individual-identity, or contains entities which only differ numerically (Section 2.4.4). I conclude the chapter with a discussion of this last result (Section 2.4.5).

2.2 Modern Interpretations of PII

The Identity of Indiscernibles (PII) is the principle that no two entities can differ *solo numero*. PII is usually understood as applying only to individuals (e.g. entities with determinate identity conditions), and is commonly interpreted as establishing a relation between the two notions of numerical identity and qualitative indiscernibility whereby qualitative indiscernibility is *sufficient* for numerical identity. Since qualitative indiscernibility is commonly defined in terms of the sharing of all qualitative properties, it is not uncommon to read that PII is the thesis that no two distinct entities can instantiate all the same qualitative properties.

2.2.1 A Brief History

Arguably, the Identity of Indiscernibles was formulated for the first time by Gottfried Wilhelm Leibniz in 1686. In his *Discourse on Metaphysics*, Leibniz states PII as the thesis that "[...] it is not true that two substances can resemble each other completely and differ only in number" (*Discourse on Metaphysics*, 1686, translated in Ariew & Garber 1989, p. 41–42).

Throughout his career, Leibniz has returned multiple times on PII, and we can find different formulations of the principle in his works. Rodriguez-Pereyra (2018, p. 49) identifies two standard formulations of PII, according to which (1) "nowhere are there things perfectly similar" (*On Nature Itself*, 1698, translated in Ariew & Garber 1989, p. 164), and (2) "it is not possible for there to be two individuals entirely alike, or differing only numerically" (*Letter to Arnauld, May 1686*, translated in Ariew & Garber 1989, p. 73). The only difference between (1) and (2) lies in their modal profile: while (1) only states that there are no entities differing *solo numero*, and is therefore compatible with a reading of it as a contingent truth, (2) states that indiscernible and yet distinct entities are *impossible*, and must be interpreted as a necessary truth. Rodriguez-Pereyra (2018, p. 49) remarks that in the twentieth century literature, many authors have regarded (2) as the true version of PII.[1] In particular, Adams (1979, p. 12–13) writes:

> Leibniz commonly states [the Identity of Indiscernibles] in the language of necessity. And well he might; for he derives [it] from his theory of the nature of an individual substance, and ultimately

[1]See, for example: Adams (1979), Parkinson (1965), Rescher (1967), and Russell (1992).

from his conception of the nature of truth, which he surely regarded as absolutely necessary.

Rodriguez-Pereyra (2018, p. 49) suggests, however, that there are reasons to think that at least in his correspondence with Clarke, Leibniz might have taken the Identity of Indiscernibles as only contingently true. He also remarks that Leibniz was concerned with *qualitative difference*, namely: difference with respect to qualitative properties. (Rodriguez-Pereyra 2018, p. 50) This means that, among the many versions of PII we will discuss in Section 2.2.2, Leibniz himself would have endorsed the one restricted to qualitative properties alone.

It is interesting to notice that in his correspondence with Clarke, Leibniz attempts to prove the Identity of Indiscernibles from his principle of Sufficient Reason, according to which every truth has an explanation.[2] The main idea behind the proof is that if there were indiscernible entities, God would still have preferred the actual world over an indiscernible world. However, God preferring one of two indiscernible worlds is a violation of the principle of Sufficient Reason, and therefore the Identity of Indiscernibles stands.

[2]Rodriguez-Pereyra (2018, p. 48) interprets the principle of Sufficient Reason as the thesis that nothing is fundamental. In his *Principles of Nature and Grace, Based on Reason*, Leibniz adopts a formulation of this principle according to which "[...] nothing takes place without sufficient reason, that is, [...] nothing happens without it being possible for someone who know enough things to give a reason sufficient to determine why it is so and not otherwise" (*Principles of Nature and Grace, Based on Reason*, 1714, translated in Ariew & Garber 1989, p. 210).

2.2.2 Many Principles

In the contemporary literature, it is common to express PII in the language of second order logic with identity as follows:

$$\forall x, y(\forall P(Px \leftrightarrow Py) \rightarrow x = y)$$

where x and y are individual variables, and P is a predicate variable ranging over properties of some specified kind.

By restricting PII's second order quantification one obtains distinct versions of the principle, whose relative strength varies along with the kind and number of properties and relations one leaves out of the quantifier's range.

French (1989, p. 144) distinguishes three versions of the principle. According to the first, which he calls PII_1, no two entities can share all their properties and relations. A stronger version, PII_2, holds that no two entities can share all their non spatio-temporal properties and relations. Finally, according to PII_3, there cannot be distinct entities sharing all their non-relational properties. Clearly, PII_1 is the weakest among the three renditions of the principle. Furthermore, since it quantifies over *all* properties and relations, it is taken by French to include also properties like 'being identical to a', which cannot be true of more than one individual, and whose inclusion makes PII just a philosophically uninteresting theorem of second order logic. Unlike PII_1, PII_2 and PII_3 are not trivial — and it is still an open question, according to French (1989), whether they hold in Classical Mechanics.

Similarly, Adams (1979, p. 11) distinguishes between (1) a trivial version of PII according to which no two distinct individuals can share all their properties, (2) a stronger version according to which no two distinct individuals can share all their qualitative properties, and (3) an even stronger version where no two distinct individuals can share all their non-relational qualitative properties.

Finally, Rodriguez-Pereyra (2006) identifies still another version of PII, according to which no two distinct individuals can share all their non-trivializing properties. According to Rodriguez-Pereyra (2006), this is the weakest definable interpretation of the Identity of Indiscernibles, and it is worthy of philosophical consideration. I challenge this version of PII in Section 2.4, where I argue that, if PII is taken as an explanation of what the identity of individuals consists in, then Rodriguez-Pereyra's version is not a viable interpretation of the principle.

Although these are so far the most discussed versions of PII (at least in Metaphysics and Philosophy of Physics), there are in principle as many versions of the Identity of Indiscernibles as there are definable collections of properties.[3] One could interpret the principle as ranging only over intrinsic properties, or physical properties, or even some gerrymandered set of properties defined by listing the relevant properties one by one. But then, what is the *correct* interpretation of PII, if there is one?

Many authors suggest that the Identity of Indiscernibles should be understood as quantifying only over qualitative properties. For example, Adams (1979, p. 11) suggests that the correct reading of PII is that "any two distinct individuals must differ in some [qualitative property], *either* relational

[3]For other versions of PII, see: Dorato & Morganti (2013, p. 594–596), French & Krause (2006, p. 10), Hoy (1984, p. 276), Quinton (1973, p. 24–25), and Swinburne (1995, p. 390–391).

or nonrelational", where qualitative properties and relations are properties and relations that can be expressed, in a language rich enough, without the use of referential items as proper names or indexical expressions.[4] On the other hand, Wiggins (2001, p. 62–63) suggests that any property and relation involving the notion of identity should be excluded by the range of PII's second order quantifier.

2.2.3 Adams and Wiggins on PII

Although Adams' and Wiggins' proposals look similar, they are indeed quite different. To appreciate the extent of this difference one could consider, for example, properties like self-indentity, loneliness and accompaniment. According to Adams (1979, p. 7), any property expressible in a sufficiently rich formal language without individual constants or other devices of singular reference to particular individuals is qualitative, and thus the correct interpretation of PII must allow the principle to quantify only over the properties and relations that can be so expressed. Self-identity seems to meet Adams' criterion, for it can be clearly expressed in the language of first order logic by using only variables and the identity symbol.[5] The same holds for properties like accompaniment and loneliness. Following Lewis (1983, p. 198), we say that an individual is accompanied whenever it exists together with some other wholly distinct individual, and lonely otherwise. Like self-identity, accompaniment and loneliness can be expressed without individual constants,

[4] This definition of qualitative properties, proposed by Adams (1979, p. 7), has recently been challenged by Cowling (2015, p. 286–287). Although I agree with Cowling, I use Adams' definition here because I find it plausible that Adams' stance on PII depends, in one way or other, on his understanding of qualitative properties.

[5] The definition of self-identity I have in mind here is: '$x = x$'.

to wit: the first order formulas '$\exists y(x \neq y)$' and '$\forall y(x = y)$' respectively. According to Adams, then, PII can safely quantify over self-identity, loneliness, or properties alike. On the contrary, Wiggins holds that including such properties in PII's range is enough to make principle circular. This is because, according to Wiggins (2001, p. 62–63):

> If we ask about the strength of the antecedent 'for all φ, $\varphi x \leftrightarrow \varphi y$' in this rendering of Leibniz's reading of its principle, then it is clear that it is as if he considers the principle to be protected from triviality by his excluding from the range of this variable 'φ' not only predicates compounded from '=' itself but also such predicates or relations as 'five miles SW of Big Ben'. Indeed, he must think all predicates presupposing place-, time- or thing-individuation are excluded. [...] Once predicables involving '=' or its congeners and its derivatives are included within the range of the variable, the formula [which expresses PII] is neither an analytical explication nor even a serviceable elucidation of identity. For the formula manifestly presupposes identity.

Wiggins seems to share Adams' concern about non-qualitative properties and relations, even if it is not clear whether they agree on the extension of such notion.[6] Issues of interpretation aside, the main point of contention between Adams and Wiggins seems to be about what Wiggins calls "predicates compounded from '=' itself" and "predicables involving '=' or its congeners and its derivatives". It should be noted here that Wiggins seems to be using

[6] Adams, for instance, doesn't seem to include properties and relations involving time-individuation among non-qualitative properties. See Adams (1979, p. 6).

the terms 'predicates' and 'predicables' interchangeably, and he seems to be using them as synonyms of 'properties'. I will follow Wiggins' terminology up to the end of the section, for a discussion of the exact meaning of these terms is clearly beyond our purposes, and it would have no impact on the clarity of the overall argument.

We have at least an intuitive understanding of the notions of "predicates compounded from '=' itself" and "predicables involving '=' or its congeners and its derivatives", for we can easily come up with predicates belonging to both sides of the divide. The predicates of loneliness and accompaniment, for example, are a clear case of predicates involving identity. On the other hand, predicates like 'being cold' and 'being a homeowner' seem to be clear cases of predicates that do not involve identity. Other cases, however, are less straightforward. Think about the predicates 'being trilateral' and 'being even'. At first sight, these predicates don't seem to involve identity. However, when we regiment them in a formal language, the relevant formulas will likely contain the identity symbol, since they will require numerical claims which would need quantification and identity to be expressed. As soon as we consider the notion of 'regimentation', however, we find ourselves dealing with the fact that any predicate can be taken as a primitive in any suitable language.

We could then try to define the notions of "predicates compounded from '=' itself" and "predicables involving '=' or its congeners and its derivatives" in a way that is similar to the one in which Adams defines the notion of qualitative properties. According to Adams (1979, p. 7), a property P is qualitative if and only if P can be expressed, "[...] in a language sufficiently rich, without the aid of such referential devices as proper names, proper adjectives and verbs [...], indexical expressions, and referential uses of definite descriptions". In the same spirit, we can understand Wiggins' notions of "predicates compounded from '=' itself" and "predicables involving '=' or its congeners and its derivatives" as those standing for properties (and relations) that *cannot* be expressed in any language lacking the resources to express or define identity — unless, of course, they are there taken as primitives. Under this reading, a tentative definition of the notion hinted at by Wiggins is the following: a predicate P is compounded from (or involves) '=' itself or its congeners and its derivatives if and only if for any language $\mathcal{L}$ and any well-formed formula φ in $\mathcal{L}$:

> Whenever P is in $\mathcal{L}$ and there is a well-formed formula φ in $\mathcal{L}$ such that $P(x)$ and φ are logically equivalent (with respect to some suitable calculus), then (1) there is a well-formed formula ψ containing the identity symbol '=' in $\mathcal{L}_=$ and such that ψ is logically equivalent to φ, and (2) there is no sub-formula of ψ which is logically equivalent to φ.

The relation between the languages $\mathcal{L}$ and $\mathcal{L}_=$ can be outlined as follows: where $\mathcal{L}$ is any language, $\mathcal{L}_=$ is the language resulting from the addition of the identity symbol '=' to $\mathcal{L}$. Clearly, if $\mathcal{L}$ already contains a symbol for identity, $\mathcal{L}$ and $\mathcal{L}_=$ are the same language. It should be noted that the above definition is not meant as implying that the only way for a language to express the identity relation is by mean of the symbol '='. This clarification is important because, much like Adams' definition of qualitative properties, our definition of "predicates compounded from '=' itself" quantifies over all possible languages.

The ideas behind the restrictions imposed in our definition is twofold. On the one hand, we must take into consideration the fact that not all predicates that *can* be expressed without the use of identity can be considered as not involving identity or any of its congeners and derivatives.

Consider, for instance, the third order binary predicate 'v', whose intended interpretation is the following: for any second order predicate 'P', the formula '$v(P, x)$' is true whenever x is P and no entity other than x is P. Informally, 'v' expresses the uniqueness of x with respect to some predicate 'P'. It seems we have good reasons to consider 'v' as one of the predicates involving identity, for 'v' could be unpacked, in a language rich enough, by means of the identity symbol. In second order logic with identity, the formula '$v(P, x)$' is equivalent to the formula '$P(x) \wedge \forall y(P(y) \leftrightarrow (x = y))$' (when the third-order predicate 'v' is correctly defined). However, the predicate 'v' doesn't need identity to be expressed.[7] Consider for example a second order language $\mathcal{L}$ in which 'v' is a primitive predicate symbol. If $\mathcal{L}$ is a language without identity, then we can clearly claim to be able to express 'uniqueness with respect to a certain predicate' in $\mathcal{L}$ without being able to express identity. This, however, shouldn't count as a reason to consider 'v' as a predicate that doesn't involve the notion of identity. This is the rationale behind condition (1) in our definition.

On the other hand, we have to deal with the fact any non identity involving formula 'φ' is equivalent to the formula '$\varphi \wedge \psi$', where 'ψ' is a tautological

[7] Or at least, it doesn't need identity at the level of the object-language to be expressed.

formula with identity. That is: given any predicate 'P' in any language $\mathcal{L}$, it is always possible to construct formulas like '$P(x) \wedge (x = x)$' in $\mathcal{L}_=$, which involve identity and are trivially equivalent to '$P(x)$'. This is the reason behind condition (2) of our definition.

The above considerations show that Wiggins' restriction on the admissible range of PII's second order quantification is wider than Adams', in the sense that the class of properties we should exclude from PII's scope according to Adams is a proper subset of the class of properties we should exclude from PII's scope according to the Wiggins.

2.3 PII & Other Criteria of Identity

The worry motivating Adams and Wiggins in advocating for such restrictions is twofold. On the one hand, they claim, PII should be protected from triviality. And were we to consider its second order quantification as unrestricted, we would most certainly end up with a logical truth, instead of a substantive metaphysical claim. For clearly, we can count among the properties that an individual x instantiates the property 'being identical to x'. Therefore, were we to have two numerically distinct individuals x and y, it would always be the case that there is a property, i.e. 'being identical to x', that only x has, and another property, i.e. 'being identical to y', that only y has. Under this reading, the Identity of Indiscernibles most certainly comes out true. This version of the principle is however rather uninteresting, since its validity is not a matter of how things are.

On the other hand, they argue, we should also protect PII from circularity. And this is the reason why, according to them, non-qualitative properties cannot be in the range of PII's second order quantifier. To explain. The Identity of Indiscernibles is a principle concerning the identity of individuals. In its bi-conditional formulation, PII gives the necessary and sufficient conditions for individuals to be numerically identical.[8] Also: in virtue of its logical form, the Identity of Indiscernibles is a *criterion of identity* for individuals — for it is a clear instance of the schema for identity criteria proposed in Lowe (1989, p. 6):

$$\forall xy((\varphi x \wedge \varphi y) \rightarrow (x = y \leftrightarrow Rxy))$$

where 'φ' is a sortal term, and R an equivalence relation.[9] Furthermore, the Identity of Indiscernibles gives us a rule for deciding, under all circumstances, whether two individuals x and y are the same. And, one could argue, the whole purpose of PII as a criterion of identity is to explain what 'grounds' the identity of individuals. (I will discuss this understanding of PII in more details in Section 2.4.) Under this reading, being non-qualitative properties somehow dependent on the identity of individuals, to include such properties in PII's range would be to explain the identity of some individuals in terms

[8] I am here referring to the following formulation of PII, which can be found in Wiggins (2001, p. 62): '$\forall\varphi(\varphi x \leftrightarrow \varphi y) \leftrightarrow (x = y)$'.

[9] In the literature about identity criteria, it is customary to distinguish between one-level identity criteria and two-level identity criteria. (See, among others, Lowe 2012 and Williamson 1990.) One-level identity criteria have the form: '$\forall xy((\varphi x \wedge \varphi y) \rightarrow (x = y \leftrightarrow Rxy))$', while two-level identity criteria have the form: $\forall xy((\varphi x \wedge \varphi y) \rightarrow (d(x) = d(y) \leftrightarrow Rxy))$. As we will see, examples of one-level identity criteria include the Axiom of Extensionality for Sets and Davidson's (1969) criterion for event identity. Examples of two-level identity criteria, on the other hand, include Frege's (1884/1950) criteria for the identity of directions and numbers. Since in this Chapter I am only concerned with one-level identity criteria, I will refer to these simply as 'identity criteria'. This will bear no consequences for my argument.

of the identity of other (possibly distinct) individuals, thus yielding a circular thesis.

It is common for identity criteria to be cast in such a way that the equivalence relation R involves the notion of identity, without this compromising their ability to explain the identity of the entities they quantify over in a non circular way. Consider for instance the identity criterion for material individuals that Horsten (2010, p. 420) refers to as 'Locke's Thesis', which claims that a material object x is the same material object as the material object y if and only if (1) x and y belong to the same kind, and (2) there is a moment in time where x and y spatially coincide. Here, R is the relation of spatial coincidence (relativised to a certain moment in time), which is intrinsically identity-involving — for x and y can spatially coincide only when they exactly occupy *the same* region of space. Nonetheless, Locke's Thesis isn't usually regarded as circular, and this is because we do not, and should not, expect it to be a definition of identity *per se*.[10] Locke's Thesis aims to explain what grounds the identity of material objects, and does it by holding that the identity of material objects is grounded on the identity of non-material ones, to wit: regions of space. And the fact that the question of identity is deferred from material entities to non-material ones does not constitute an argument against Locke's Thesis.

Now, one might ask whether the Identity of Indiscernibles is at all similar to Locke's Thesis, for the first order quantifier in PII seems, at first sight, unrestricted. I hold that, exactly as for Locke's Thesis, it would be mistaken

[10] This doesn't mean that Locke's Thesis is a good criterion of identity. The criterion has been in fact challenged on independent grounds, among others, by Fine (2000) and De Clercq (2005).

to believe that the Identity of Indiscernibles, in any of its restricted or unrestricted interpretations, could ever be understood as a definition of identity *per se*, as opposed to a statement about what the identity of a chosen class of entities, however large, consists in. The Identity of Indiscernibles is not a definition of identity, and it cannot be: for irrespective of any restriction on its quantification, the relation R at the right-hand side of PII's bi-conditional does indeed involve identity, and it does so necessarily. The principle holds in fact that any two entities x and y are the same whenever they instantiate all the *same* properties and relations (of a given kind). In a famous passage about the Axiom of Reflexivity of Identity and the Indiscernibility of Identicals, Lowe (2016, p. 52) claims:

> I don't regard [the Axiom of Reflexivity of Identity and the Indiscernibility of Identicals] as providing even an implicit *definition* of identity [...]. No one could *learn the meaning* of the term "identity" from grasping these axioms, because an understanding of identity is already required in order to grasp them (for instance, it must be grasped that "φ" in its two different — *nonidentical* — occurrences in [the Indiscernibility of Identicals] should always be given the *same* interpretation).

I suggest that Lowe's words can be applied, *mutatis mutandis*, to PII, and that therefore it would be wrong to consider PII a definition of identity *per se*.

2.3.1 Strong and Weak Circularity

There is however one important difference between the Identity of Indiscernibles and Locke's Thesis. We saw that Locke's Thesis explains the identity of material objects in terms of the identity of spatial regions. And since spatial regions are not material objects themselves, then charging Locke's Thesis with circularity would be, at the very least, unfair. That is: Locke's Thesis avoids the circularity charge by appealing to the difference *in kind* between the objects whose identity is allegedly explained and the objects whose identity is used as an explanation. The same however wouldn't hold for PII, were we to include non-qualitative properties in the range of its second order quantifier. For clearly in such case the principle would ground the identity of (at least) some individuals on the identity of other (possibly distinct) individuals, thus falling into the alleged circularity.

In this respect, PII is more similar to some other well-know identity criteria: the Axiom of Extensionality for sets, and Davidson's (1969) criterion of identity for events. The Axiom of Extensionality (AE) holds that sets with the same members are the same set. Formally: $\forall xy((x = y) \leftrightarrow \forall z(z \in x \leftrightarrow z \in y))$ — where x, y and z are sets.[11] On the other hand, Davidson's criterion of identity for events (EI) holds that events with the same causes and effects are the same event — where only events can be causes, as well as effects, of other events.[12] The Axiom of Extensionality and Davidson's

[11] Here, the Axiom of Extensionality is understood as quantifying only over sets. In this sense, AE is a criterion of identity for pure sets, i.e. those sets that can only have sets as members. Although the axiom can be generalized as to cover impure sets, i.e. those sets which admits non set-like entities as members, the version that interests us here is the restricted version, as it appears in Zermelo-Fraenkel's Axiomatic Theory of Sets ZF(C).

[12] Formally: $(x = y) \leftrightarrow \forall z((z \text{ causes } x \rightarrow z \text{ causes } y) \wedge (x \text{ causes } z \rightarrow y \text{ causes } z))$,

criterion of identity for events are alike in one important respect: they understand the identity of some individuals of a given kind (respectively: sets and events) as depending on the identity of entities of the *same* kind. In this respect, they are more similar to an unrestricted PII than they are to Locke's Thesis. And one might worry whether, as result of this similarity, AE and EI face the same charge of circularity as PII.

In the debate on identity criteria, circularity is usually seen as a consequence of impredicativity.[13] Lowe (1989, p. 178) distinguishes two senses in which a criterion of identity can be said to be impredicative. In particular, a criterion of identity is weakly impredicative whenever it quantifies over "[...] a totality which includes the very entities for which it supplies a criterion of identity". In this sense, Lowe remarks, all identity criteria of the form '$\forall xy((\varphi x \wedge \varphi y) \rightarrow (x = y \leftrightarrow Rxy))$' are weakly impredicative. On the other hand, a criterion of identity is strongly impredicative whenever it features, *on the right-hand side of the bi-conditional*, "[...] a quantifier binding variables which range over a totality which includes or depend on the very entities for which an identity criterion is being supplied".

Following Lowe's distinction, we can say that PII, AE and EI are impredicative in both the weak and the strong sense. They are weakly impredicative in virtue of their logical form, and strongly impredicative for they quantify, at the right-hand side of their respective bi-conditionals, over totalities which include the entities whose identity is at issue. In the case of AE and EI, strong impredicativity is apparent in the surface logical grammar: they range over

where x, y and z are events. This formulation of the principle is found in Davidson (1969, p. 231).

[13]See, among others, Horsten (2010) and Quine (1985).

the totality of sets and events respectively, and those totalities include the entities whose identity is being defined. On the other hand, identifying the strong impredicativity that affects the unrestricted interpretation of PII requires a more comprehensive analysis. Prima facie, in fact, the principle ranges over the totality of properties, and properties are not individuals.

However, some of the properties the unrestricted principle quantifies over have concrete constituents as per Rosenkrantz (1979), and therefore depend on individuals. Consider, for instance, the property 'being five feet away from Donald Trump' and the property 'being five feet away from Joe Biden'. The identity of these two properties (call them P and Q respectively) is entirely dependent upon the identity of Donald Trump and Joe Biden. With this, I mean that the the fact that P and Q are distinct properties is a consequence of the fact that Donald Trump and Joe Biden are distinct individuals. Furthermore, the possibility that some entity instantiates P and Q will itself depend on certain identity facts about Donald Trump and Joe Biden. No entity could in fact instantiate P and Q in a world devoid of both Trump and Biden. This can help us make sense of Adams' and Wiggins' suggestions from another point of view. If it is true that circularity is a result of impredicativity, then by eliminating impredicativity we should expect the circularity to go away — and this is exactly what Adams and Wiggins suggest, namely: to eliminate from the scope of PII those properties and relations that make for the strong impredicativity of the principle.

However, *contra* Quine (1985), Lowe (1989) argues that impredicative identity criteria, even strongly impredicative ones, do not automatically display circularity in virtue of their impredicativity. Lowe (1989) considers both

AE and EI, and argues that only the second is indeed circular. In Zermelo-Fraenkel's Axiomatic Theory of Sets, in fact, (1) there is (at least) one set which has no members, and (2) for any set x, either x is a set obtained by repeated applications of the power-set operation on the empty-set, or there is a set y such that x is a subset of y and y is obtained by repeated applications of the power-set operation on the empty-set. Together with (1), AE guarantees that there is only *one* empty set. And this fact, together with (2), entails that "[...] any identity question concerning sets can ultimately be settled by reference to the empty set through repeated applications of [AE]" (Lowe 1989, p. 180). On the contrary, EI is not supported by any theory with a framework comparable to that of ZF(C), and there seems to be no *prima facie* reason for it to be. As a consequence, EI is subject to potential circularity, for it "[...] provides no effective way of determining whether an event whose identity is presupposed in fixing the identity of another [...] is in fact distinct from that other" (Lowe 1989, p. 181).

Horsten (2010) too argues that impredicativity alone doesn't necessarily undermine the degree of acceptability of an identity criterion. A crucial gambit in Horsten's argument is to defend, with Williamson (1990) and Lowe (1998), a notion of identity criteria according to which they should be regarded as *metaphysical principles*, as opposed to semantical or epistemic principles. According to Horsten (2010, p. 415–419), identity criteria are not about sameness of reference, nor they contain any (basic) epistemic component. Identity criteria are metaphysically necessary theses, involving metaphysical vocabulary, and express the necessary and sufficient conditions for the identity of individuals (or more generally: entities) of some given kind. And once identity criteria are accepted as such, Horsten (2010, p. 425) claims, the argument from impredicativity to circularity to unacceptability does not follow through anymore. In particular, Horsten suggests that the charge of circularity against Davidson's principle is unfair. Sure, as it stands EI is not backed up by any theory whose framework is akin to that of ZF(C). However, EI is a metaphysical principle, and as such it imposes certain restrictions on reality by defining a class of suitable possible causal structures for it. In other words, EI's circularity (and the extent of it) depends entirely upon the structure the principle is supposed to be applied to.

Following Horsten I distinguish two separate readings of PII, which exhibit distinct kinds of circularity. Suppose we have individuals a and b, and we want to decide whether they are the same individual or not. Suppose further that we have no way to determine whether a and b are identical or not by finding some qualitative property or relation that can discern them — that is, a and b are qualitatively indiscernible. Clearly, were we to hold, as per Adams (1979) and Wiggins (2001), that PII must be restricted to qualitative properties, we would conclude that, in the given scenario, a and b are indeed one and the same individual. (Provided we would indeed *endorse* such version of PII.) What would happen however, if we decided to employ, *contra* Adams and Wiggins, an unrestricted version of the principle? We would, I submit, fall into one of two distinct cases. In the first case, we would try to tell a and b apart by appealing to properties that depend on a and b themselves. Examples might be the properties of 'being (identical to) a' and 'being (identical to) b', which only a and b can respectively instantiate, as well as properties like 'being at some non null distance from a', which only

b (among a and b) can instantiate, provided some conditions obtain in the background (i.e. a and b are not co-located objects, and so on). In such case, our application of the Identity of Indiscernibles would be inherently circular. No matter how reality is, appealing to properties depending on the identity of a and b would make our decision on the identity of a and b dependent on the identity of a and b. In other words, the principle would decide whether a and b are the same individual only after having settled the identity of a and b.

In the second case, we would try to tell a and b apart by appealing to properties that depend on the identity of other individuals, for example c and d. We would then have to decide whether c and d are the same individuals or not, and we might at that point appeal to the identity of further entities e and f. And depending on the metaphysical structure in which these individuals are considered, we might end up in one of three distinct scenarios. In the first scenario, we reach some entities g and h whose identity or distinctness can be decided only on the basis of the identity or distinctness of some individuals whose identity we had decided in some previous step. In this case, our application of the principle would indeed be circular, exactly as the application described in the previous paragraph. However, we might find ourselves in a second scenario, where the identity of any pair of individuals can be decided on the basis of the identity of further individuals, *ad infinitum*. In such case, we would clearly end up with an infinite regress. However, this regress would not result in any circularity. Finally, there is a third scenario. Similarly to scenario number two, we go on and on deciding on the identity of some individuals by resorting to the identity of other individuals. This time, however, the process is not infinite. We reach a last step, where (1) either the identity of the individuals at that step can be settled by appealing to their qualitative differences, or (2) their identity is a brute fact, or (3) the entities we meet at that last step are not individuals, and we can decide their identity by means of a different principle than PII.[14]

This shows that a reading of PII that doesn't restrict the second order quantification of the principle to qualitative properties alone can exhibit two distinct kinds of circularity: a strong circularity, and a weak one. The strong circularity is a result of letting the principle quantify over properties that depend on the identity of the exact individuals the principle is supposed to be applied to. This kind of circularity is indeed dangerous. However, it can be avoided with a very narrow restriction: to let out of the range of the principle's second order variables all the properties that either (1) depend on the identity of x and y, or (2) depend on the identity of any individual z whose identity depends, in turn, either on the identity of x or on the identity of y. This restriction is clearly narrower than both Adams' and Wiggins' suggestion. We don't need to exclude all the non-qualitative properties there might be, only selected ones. (This is, as far as I can see, the idea behind Rodriguez-Pereyra's (2006) understanding of PII.) Exactly as per Davidson's criterion, a PII so restricted would be circular only if applied to certain structures — and this might not be enough for rejecting such reading of the principle, if the right way to interpret it is as a metaphysical (or theoretical) principle, rather than a semantic or epistemic one.[15]

[14] I spell out a similar argument in Section 2.4.4.

[15] For more on this discussion, see Horsten (2010).

2.3.2 Is A Circular Identity Criterion Unacceptable?

Now that the extent of the circularity charge against an unrestricted version of the Identity of Indiscernibles has been clarified, a methodological question is incumbent upon us whether a principle displaying such circularity is indeed to be disqualified as unsatisfactory without further investigation. In defending EI against Quine's circularity charge, Horsten (2010, p. 425) suggests that Gödel's famous account of impredicative definitions might be applied *mutatis mutandis* to identity criteria:

> If [Gödel's account] is correct and criteria of identity are conceived as metaphysical principles [...], then it would seem that Gödel's remarks should also apply to identity criteria. There would appear to be nothing in the least absurd in the existence of kinds of objects for which only circular identity criteria exist. If a criterion would somehow create the objects of the sort under investigation, then circular criteria would be problematic. But since the objects exist independently of the criterion, some criteria may well *have to be* circular.[16]

And individuals, like events, may well be some of the entities for which any identity criterion which aims at full generality has to exhibit some sort of circularity. And this, I submit, might be true of individuals peculiarly, in

[16]Horsten is referring here to Gödel (1946, p. 127–128): "[...] it seems that the vicious circle principle [...] only applies if the entity involved are constructed by ourselves. In this case there must clearly exist a definition [...] which does not refer to a totality of things to which the thing to be constructed itself belongs. If, however, it is a question of objects that exist independently of our constructions, there is nothing in the least absurd in the existence of totalities which can be described [...] only by reference to this totality".

that in our understanding of individuals we seem not to be able to dispense with a certain pre-theoretical understanding of identity itself. For what is an individual if not something which is identical only to itself and distinct from any other individuals? I hold that our understanding of individuals is so entrenched with our pre-theoretical understanding of identity, that it is hard to see how any of the two notions would survive unharmed, were the other to be gone. There would appear to be nothing in the least absurd then, if any criterion of identity for individuals *had to be* somehow circular. And individuals, as well as events, can be consistently conceived (and indeed they are so conceived by the vast majority of philosophers) as mind-independent entities, enjoying a separate existence from any of our definitions.

Furthermore, a circular criterion (at least a weakly circular one, in the sense of weak circularity explained above) does not automatically become uninformative. This fact, which Horsten (2010) shows with respect to EI, holds equally, I suggest, for PII. Informativeness is listed by Horsten (2010, p. 422) among the adequacy conditions of any criterion of identity, and it is defined as the ability of the criterion to "[...] impose ontological constraints to the non-logical relations and properties in terms of which it is formulated". In this sense, Horsten continues, "[...] informativeness comes in degrees. If one identity criterion excludes more models than another identity criterion, then the former is more informative than the latter". And we just saw that a weakly circular PII, unlike the strongly circular one, succeeds in imposing such ontological constraints. In other words: the fact that a weakly circular PII is informative while a strongly circular one is not can be explained by seeing that, while the second is a truth of logic, the first is not.

Finally, it could well be argued on behalf of a weakly circular PII, that apart from giving us an account of what the identity of individuals consists in, it embeds some information about the very notions of 'identity' and 'individuality', namely: that these two notions are somehow interdefinable, and although occupying distinct places in our overall ideology, no one of them can survive without the other.

2.4 PII's Weakest Interpretation

However, you might hold for independent reasons that identity and individuality are not so intertwined as a weakly circular version of PII would demand, and you might also believe that identity is not fundamental. Indeed, you might hold, in the good company of Adams (1979), French (1989), Wiggins (2001), Della Rocca (2005), and Hawley (2009), that PII should offer a way to *explain*, or *define*, the concept of individual-identity (i.e. identity between individuals). On other words: you might hold that the best PII has to offer is the possibility to dispense with individual-identity as a primitive concept in our theorising, and that there is no reason for including a principle as strong as PII in our theories, if in return we cannot explain the identity between individuals in some more fundamental terms.

This notion of *explanation*, or *definition*, can be better understood with an example. Consider again the Axiom of Extensionality, according to which sets x and y are identical if and only if they have the same members. This axiom can be considered as an explanation of what the identity of sets consists in, and a definition of the concept of identity within Zermelo-Fraenkel's Axiomatic Set Theory (ZFC). It explains the concept of set-identity in the sense that it explains identity facts about sets in terms of facts about membership. It also defines the concept of set-identity since it give us a way to dispense with identity as a primitive notion: in ZFC, the only primitive relation is membership, and this is sufficient to settle any identity facts pertaining to sets.

More generally we can say that a concept C is explainable/definable whenever, given a formula φ, (1) we have a systematic way of replacing any subformula of φ in which C appears with a formula in which C doesn't appear, and (2) the result of this substitution preserves both the truth and the intended meaning of φ. In ZFC, as we have seen, we can substitute any formula in which the identity symbol appears with a formula where only the symbol for membership appears. Analogously, if Adams' interpretation of PII is true, then we have a systematic way to replace any formula containing the identity symbol flanked by individual constants or variable with a formula where no identity symbol appears.

It is important to keep in mind that even if we can in principle explain identity in terms of other notions, this doesn't tell us anything about the ontological status of identity — for as we all know there are good and bad explanations, and philosophical arguments are often required, to the effect that a certain explanation is indeed better than another one. A very famous example is that of the concept 'grue'.[17] We can explain the concept of 'grue' by means of the concepts of 'green' and 'blue'. But we can also go the other way around, and explain the concept of 'green', and the concept of 'blue',

[17] See Goodman (1965, p. 74).

by means of the concept of 'grue'. Which way provides us with the best explanation is ultimately a matter for philosophical debate.

However, if you believe PII should offer an explanation of individual-identity in these terms, then even a weakly circular version of PII would be off the table. In this Section I argue that the weakest viable interpretation of PII available to those who seek to explain this kind of identity in terms of other more fundamental notions is the one proposed by Adams (1979) according to which, if individuals x and y share all their qualitative properties, then they are identical.

2.4.1 A Non-trivial Version of PII

In his seminal *Individuals: An Essay in Descriptive Metaphysics* Strawson (1959, p. 120) writes:

> [...] in the only form in which it is worth discussing, [the principle of the Identity of Indiscernibles claims that] it is necessarily true that there exists, for every individual, some description [involving only qualitative properties], such that only that individual answers to that description.

Rodriguez-Pereyra (2006) challenges Strawson's intuition and introduces a weaker reading of PII according to which no two individuals can share all their non-trivializing properties. This new reading of PII, he claims, can be maintained whilst rejecting Strawson's stronger version.

The difference between the readings proposed by Strawson (1959) and Rodriguez-Pereyra (2006) lies in how the principle's second order quantification is restricted. Rodriguez-Pereyra's reading is weaker than Strawson's since, as we will see shortly, the class of non-trivializing properties is taken to properly include the class of qualitative properties.

2.4.2 Trivializing Properties

What are non-trivializing properties? According to Rodriguez-Pereyra (2006, p. 205–206), *trivializing* properties are those properties whose inclusion in PII's second order quantification makes the principle a theorem of second order logic. Striking examples of trivializing properties are *identity properties*, like 'being identical to Joe Biden'. As we have seen in Section 2.2.2, including identity properties in the range of PII's quantification trivializes the principle. Other examples of trivializing properties are: 'being numerically distinct from Joe Biden', 'being identical to Joe Biden or being tall', and 'being a member of Joe Biden's singleton'.

Among the non-trivializing properties, Rodriguez-Pereyra continues, we find all the usual qualitative properties (like 'being tall', 'being wise', etc.) as well as some non-qualitative properties, like the property 'standing next to Joe Biden'. The distinction between qualitative and non-qualitative properties, although being intuitively clear, is famously difficult to make theoretically precise.[18] Informally, we say that a property P is qualitative whenever it doesn't depend on the identity of any individual, and non-qualitative otherwise. Rodriguez-Pereyra (2006, p. 205) explicitly defines non-qualitative properties as those depending on the identity of some *relatum*.

[18]The literature on the qualitative/non-qualitative distinction is vast. See, among others: Cowling (2015, 2021), Eddon (2010), Hoffmann-Kolss (2019), and D. Locke (2012). I will attempt a precise charaterisation of qualitative properties in Chapter 5.

From this definition, it follows that the general form of any simple non-qualitative property is 'standing in R to i_1 ... i_n', for R some relation and $i_1 ... i_n$ some specific individuals. For simplicity, in the remainder I will take the general form of a non-qualitative property to be: 'standing in R to i'. This will not reduce the scope nor the strength of the argument I present in Section 2.4.4, for it can be easily extended to cover simple properties with multiple relata as well as complex properties, that is: properties obtained by combining simpler properties through logical operations.

Since not all non-trivializing properties are qualitative, Rodriguez-Pereyra's interpretation of PII quantifies over (at least) some non-qualitative property. It is in this sense that his reading of PII is weaker than Strawson's: by considering a larger set of properties, Rodriguez-Pereyra's reading recognises as (metaphysically) possible many situations that Strawson's one would dismiss as impossible.

2.4.3 Setting the Stage

In what follows I label Strawson's and Rodriguez-Pereyra's readings of the Identity of Indiscernibles PII-S and PII-RP, respectively. I will thus refer to the thesis that no two individuals can share all their qualitative properties as PII-S, and to the thesis that no two individuals can share all their non-trivializing properties as PII-RP. I argue that Rodriguez-Pereyra's PII-RP is incompatible with the claim that the Identity of Indiscernibles should explain individual-identity. If I am correct, it follows that those who find themselves in the good company of Hawley (2009) and Wiggins (2001) will not find

PII-RP attractive. On the other hand, those upholding PII-RP will have one extra reason to judge this kind of identity as not amenable to further explanation.

To this conclusion, I will argue that PII-RP, taken in isolation, leads to an infinite regress in the explanation of individual-identity, avoided only at the cost of admitting the possibility of individuals differing *solo numero.* And if one is seeking for an explanation of individual-identity, none of these options will be adequate. Here is why. With Rodriguez-Pereyra (2018, p. 50), we say that individuals a and b differ *solo numero* whenever "[...] their difference is simply due to the fact that one of them is a and the other is b". If this is correct then, if indeed there are some a and b whose distinctness is a case of *solo numero* difference, then there is no identity-free formula that we can substitute 'a is distinct from b' with, while preserving truth and intended meaning. *Solo numero* difference leads inevitably to cases where identity cannot be explained. The same is true in cases where the explanation of identity sets off an infinite regress: if the distinctness of a and b depends on the identity of c, and the distinctness of a, b, and c depends on the identity of d (and so on), the formula 'a is distinct from b' cannot be substituted with any identity-free formula with the same intended meaning.

2.4.4 An Argument Against PII-RP

Let S be a situation where PII-RP is true and PII-S false, and all entities in S are individuals. By definition, in S: (1) there are two individuals a and b which share all their qualitative properties, and (2) a and b differ over some non-trivializing, non-qualitative properties.

As per (2), let P be one such non-trivializing non-qualitative property. Because P is non-qualitative, it is of the form 'standing in R to i' for R some relation and i some specific individual. And because P discerns between a and b, then either $P(a)$ or $P(b)$, but not both. Let us assume without loss of generality that $P(a)$. We have now two cases:

> **CASE 1:** either i is a or i is b.
>
> **CASE 2:** i is a different individual c.

CASE 1: Suppose i is a. Then P has the form 'standing in R to a'. Since $P(a)$ then also $P_1(a)$, where P_1 is a property of the form 'standing in R to some entity x such that $Q_1(x)$, $Q_2(x)$, ...', and Q_1, Q_2, etc. are all of a's qualitative properties.[19] Orthodoxy has it that conjunctions of qualitative properties are themselves qualitative properties. (See Adams (1979) for a convincing defense of this idea.) A fortiori, P_1 is a qualitative property, since it is just a conjunction of qualitative properties.

As per (1), a and b are qualitatively indiscernible; it follows that $P_1(b)$ is true, with Q_1, Q_2, etc. being all of b's qualitative properties. Since P_1 makes reference to an object x[20], we should now ask: which object is x? Again we have three cases:

[19]Conjunctive properties like P_1, of perhaps infinite complexity, are not new to the literature. Compare with Adams (1979), Hoffmann-Kolss (2019), and Strawson (1959).

[20]I say that a property P makes reference to some object x whenever P is of the form 'standing in R to x'. If you think of properties as structured entities, then P makes reference to x whenever P includes x.

CASE 1.1: x is a. Then $P(b)$ is true, contradicting our initial assumption.

CASE 1.2: x is b. Then $P_2(b)$ is true, where P_2 is a property of the form 'standing in R to b'. As I see it, this option divides in two cases: either P and P_2 are the *only* properties discerning between a and b, or not.

CASE 1.2.1: Assume P and P_2 are the only properties discerning between a and b. I claim a and b differ *solo numero*. This follows from Rodriguez-Pereyra's definition of non-qualitative properties, recall, properties whose instantiation depends on the identity of specific objects. As a result, if P and P_2 are the only properties discriminating a from b, the distinctness of a and b depends on the instantiation of P and P_2; yet, the instantiation of P and P_2 depends on the identity of a and b; hence, the distinctness of a and b circularly depends on the identity of a and b.

CASE 1.2.2: Assume there are other properties discerning between a and b. Since PII-S fails in S, these other properties are non-qualitative. We have two cases:

CASE 1.2.2.1: All these other properties only make reference to a or b. Then, by an argument parallel to **CASE 1.2.1**, a and b differ *solo numero*.

CASE 1.2.2.2: There is some non-qualitative property that discerns between a and b and makes reference to some individual different from a and b. Let P_3 be some such property. Now, P_3 is of the form 'standing in R to i', for i different from a and b. Therefore there is some extra individual c in S such that i is c. And because P_3 discerns between a and b, then either $P_3(a)$ or $P_3(b)$.

Assume without loss of generality that $P_3(a)$ is true. Then a has the property 'standing in R to c'. Similar to before, let U_1, U_2, ... be all of c's qualitative properties. Then $P_4(a)$, where P_4 is a property of the form 'standing in R to some x such that $U_1(x)$, $U_2(x)$, ...'. Since a and b are qualitatively indiscernible, then also $P_4(b)$.

If the x referred to in P_4 is c, then $P_3(b)$ contradicting our earlier assumption that P_3 is discerning. Suppose then that x is a; it follows that $U_1(a)$, $U_2(a)$, etc., and because a is qualitatively indiscernible from b, also $U_1(b)$, $U_2(b)$, etc. A parallel reasoning applies if x is b. From this we conclude that if the individual to which b is related according to P_4 is either a or b, then S contains *three qualitatively indiscernible* individuals: a, b, and c. And since by assumptions *only* PII-RP holds, then: (3) there is some non-qualitative property P_5 discerning between b and c, and (4) there is some non-qualitative property P_6 discerning between a and c.

Being non-qualitative, P_5 is a property of the form: 'standing in R to j' for j some specific individual. It should be clear that if P_5 is the only property discerning between b and c, and j is either b or c then b and c differ *solo numero*. If j is either b or c, and P_5 is not the only property discerning between b and c but all the other properties which discern them only make reference to b or c, then again b and c differ *solo numero*. Finally, if j is either b or c, and there are other properties discerning between b and c and they make reference to individuals different from b and c then:

either all of them make reference to a, in which case the distinctness of b and c depends on the identity of a which in turn depends on the identity of c (and therefore we have a circularity and individual-identity is unexplainable), or they make reference to some other individual d, at which point we just reiterate the same argument, setting off an infinite regress in the explanation of individual-identity.

Suppose then that j is a. Again, if P_5 is the only property discerning between b and c, the distinctness of b and c depends on the identity of a which in turn depends on the identity of c (and therefore we have a circularity and individual-identity is unexplainable). If P_5 is not the only property discerning between b and c but all the other properties which discern them only make reference to a or b or c then we have the same circularity. Finally, if there are other properties discerning between b and c and they make reference to individuals different than a, b, and c then they must make reference to some other individual d, at which point we just reiterate the argument again.

Then j must be yet another entity, d, which is qualitatively identical to a, b, and c, and yet numerically distinct from all of them. And, as you can see, we are back again at the starting point, this time with a new individual to consider, and some other non-qualitative properties discerning between all individuals in S. (The same holds for P_6.)

Since this reasoning applies to any property other than P and P_2 which discerns between a and b without making reference to any of them, we conclude that: if the entity x in P_1 is either a or b, then either we find a contradiction, or there are entities which differ *solo numero*, or there are infinitely many entities, such that their identities depend on one another. Neither option is viable if we want PII to explain individual-identity.

CASE 1.3: Assume the x referred to in P_1 is some other individual, c, distinct from both a and b. Then $Q_1(c)$, $Q_2(c)$, etc., are true. Observe a, b and c are qualitatively indiscernible yet numerically distinct. So either we have a case of difference *solo numero*, or there are some non-qualitative non-trivializing properties P_7 and P_8 which discern respectively between a and c (P_7), and between b and c (P_8).

Assume without loss of generality that $P_7(a)$ is true while $P_7(c)$ is false. Since P_7 is non-qualitative it is a property of the form 'standing in R to i'. Suppose i is a; then a has the property of 'standing in R to a'. Hence, a falls under the property P_9 of 'standing in R to some x such that $Q_1(x)$, $Q_2(x)$, ...', where Q_1, Q_2, etc., are again all of a's qualitative properties. But then also $P_9(c)$ is true, for P_9 is qualitative and a is qualitatively indiscernible from c.

Yet if the x referred to in P_9 is a it follows that $P_7(c)$, contradicting our assumption. If x is b, then the distinctness of a and c depends on the identities of a and b. However, the distinctness of a and b depends on the identity of c, meaning that the identity of either a, b, or c is primitive. Now if x is c, then, as before, either a and c differ *solo numero* (in case P_7 is the *only* property discerning them or all the properties that discern them only make reference to either a or c) or we have an infinite regress (in case there are also other properties that make reference to entities other than a, b, or c). If any of those properties makes reference to b then again the distinctness of a and c is dependent on the identities of a and b. The same reasoning applies to i in P_7 being either b or c.

Finally, as a result, the x referred to in P_9 is some distinct entity d, qualitatively indiscernible from a, b and c. The same holds for b, c and P_8.

This regress only stops on pain of conceding that some of the individuals in S differ *solo numero*. Hence **CASE 1** entails the following: any situation in which only PII-RP is true either admits *solo numero* difference, or contains infinitely many individuals and engenders an infinite regress in the explanation of their identities.

CASE 2: The individual i referred to in P is distinct from both a and b. (Recall, P is of the form 'standing in R to i', and that $P(a)$ is true while $P(b)$ is false.) Call this individual c. Then P is of the form 'standing in R to c'.

As usual, let U_1, U_2, etc., be all of c's qualitative properties. Then $P_{10}(a)$ is true, where P_{10} is a property of the form 'standing in R to some x such that $U_1(x)$, $U_2(x)$, ...'. Since a and b are qualitatively indiscernible, $P_{10}(b)$ is true.

If the x referred to in P_{10} is c then $P(b)$ is true, contradicting our assumption. Suppose x is a; it follows that $U_1(a)$, $U_2(a)$ ..., and $U_1(b)$, $U_2(b)$ ..., by qualitative indiscernibility. (Similarly if x is b.) Hence, if the individual to which b is related according to P_{10} is either a or b, then S contains *three qualitatively indiscernible* individuals: a, b, and c. And since only PII-RP is true in S, we have the following: (5) there is some non-qualitative property P_{11} discerning between b and c, and (6) there is some non-qualitative property P_{12} discerning between a and c.

We have now only to repeat our familiar argument. Since P_{11} is non-qualitative it is a property of the form: 'standing in R to j', where j is a specific individual. If P_{11} is the only property discerning between b and c and j is either b or c, then b and c differ *solo numero*. (Same reasoning again.) If j is either b or c and P_{11} is not the only property discerning between b from c but all the other properties which discern them only make reference to b or c, then again b and c differ *solo numero*. Finally, if j is either b or c and there are other properties discerns between b and c and they make reference to individuals other than b and c, then either all of them make reference to a, in which case the distinctness of b and c depends on the identity of a which in turn depends on the identity of c (and therefore we have a circularity and individual-identity is unexplainable), or they make reference to some other individual d, at which point we reiterate the argument again.

Assume then that j is a. Again if P_{11} is the only property discerning between b and c, then the distinctness of b and c depends on the identity of a which in turn depends on the identity of c. If P_{11} is not the only property discerning between b and c but all the other properties which discern them only make reference to a or b or c, then again we have a circularity. Finally, if there are other properties discerning between b and c and they make reference to individuals other than a, b and c, then they make reference to some other individual d, at which point we just reiterate the argument.

Then j is yet another entity, d, which is qualitatively indiscernible from a, b, and c, though numerically distinct from all of them. It should be clear that we are back again at our starting point, this time with a new individual to consider, and some other non-qualitative properties discerning between all individuals in situation S. (The same holds for P_{12}.)

The argument shows that any scenario in which only PII-RP is true, is either a scenario which admits of *solo numero* difference, or a scenario that contains infinitely many individuals, and where the explanation of the relevant identities constitutes an infinite regress.

2.4.5 Philosophical Remarks

Rodriguez-Pereyra's reading of the Identity of Indiscernibles entails an infinite regress in the explanation of individual-identity, avoided only at the cost of accepting individuals which differ *solo numero*. It follows that PII-RP alone cannot be an adequate reading of the Identity of Indiscernibles if the principle is supposed to allow us to explain this kind of identity.

As we have seen, *solo numero* difference leads to cases where individual-identity cannot be explained. If individuals a and b differ only numerically, then their difference is simply due to the fact that one of them is a, and the other is b. No other facts can be found that account for their distinctness: if a and b differ *solo numero*, then they are distinct entities just *because* they are not the same entity. The identity facts concerning a and b are therefore resting on nothing else. And if this is this case, then there are some cases in which individual-identity is just fundamental, or brute.

In an infinite regress like the one PII-RP ensues, the distinctness of any two entities x and y can only be explained by identity facts about some other entity z. Someone might hold that this kind of regress is not vicious. (See

Section 2.3.1.) After all, in all situations in which only PII-RP is true we have a way to explain the distinctness of any two entities: all we have to do is to include in the relevant explanation the identity of some other entity. To those seeking an explanation of individual-identity, however, this is as much unsatisfactory as it is correct. (For discussion on these types of regress, and whether they are vicious, see Cameron 2022, Chapter 1 and the references therein.) Although we can in fact explain the identity of any two individuals in terms of the identity of an other individual, we lack the kind of generality that people looking to explain individual-identity call for. Is it not enough that we can explain *this* identity fact and *that* identity fact: if we don't have a way to explain *all* identity facts without mentioning further identity facts, we have not explained individual-identity. Also, for any identity fact that we can explain, the notion of identity appears in our explanandum as well as in our explanans: and this is not an explanation in the sense discussed above.

There is also a further problem with PII-RP, which concerns the status of the Identity of Indiscernibles as a criterion of identity for individuals, for a criterion of identity should not entail facts about the number of possible entities. And this is exactly what PII-RP does. For if we hold that *solo numero* difference is unacceptable, then PII-RP commits us to an ontology with denumerably infinitely many entities — and this can't be right. However we want to look at it, a thesis like PII-RP should not exclude that there are only finitely many entities.

For all these reasons, I hold Rodriguez-Pereyra's reading of PII is incompatible with the view that the Identity of Indiscernibles should provide an explanation of individual-identity. My argument against PII-RP generalises to any interpretation of the Identity of Indiscernibles that is *strictly* weaker than Strawson's PII-S. This is because it hinges solely on the fact that PII-RP quantifies over some non-qualitative properties, a feature shared by any reading of the principle that is weaker than PII-S: no further assumption is required for the argument to go through. In light of this, I conclude that the weakest interpretation of the Identity of Indiscernibles which can be of any use in explaining individual-identity is one quantifying only over qualitative properties.

Chapter 3

Against the Identity of Indiscernibles

3.1 Introduction

In this Chapter I take forward the discussion of the Identity of Indscernibles I started in Chapter 2. Here I focus on the question of whether a version of PII which is restricted to qualitative properties should be accepted as a necessary truth or not. This Chapter is divided into three main Sections.

In Section 3.2 I discuss three counterexamples to PII, put forward by Kant (Section 3.2.1), Adams (Section 3.2.2), and Wüthrich (Section 3.2.3) respectively. Together with the counterexamples to PII already discussed in Section 1.1.1, they complete the collection of the most influential challenges to PII within Metaphysics and Ontology.

In Section 3.3 I present and discuss the most common strategies that have historically been employed to defend PII from alleged counterexamples. These are the so-called identity defense (Section 3.3.1), the so-called discerning defense (Section 3.3.2), the so-called summing defense (Section 3.3.3), and what I call the structure defense (Section 3.3.4). Common to all four strategies is one key idea: that no counterexample to PII can be really deemed successful, which could be alternatively described, keeping all qualitative aspects of its original description unchanged, as to be compatible with the Identity of Indiscernibles. This gives justice to Hacking's (1975, p. 255) intuition that PII is ultimately a "metaprinciple about possible descriptions".

Having seen how to shield PII from possible challenges, in Section 3.4 I present a new counterexample to the Identity of Indiscernibles using branching worlds: i.e. worlds with multiple incompatible time-lines diverging as a consequence of indeterministic events. I present the counterexample in Section 3.4.1, and I argue for its possibility in Section 3.4.2. After responding to an objection according to which the branching world I set up runs against common intuitions about identity and persistence (Section 3.4.3), I argue that, unlike Black's two-spheres world, this new counterexample is successful against *all* lines of defense of PII discussed in Section 3.3. This is, to me, its most interesting feature, for a quick look at the relevant literature reveals that many authors dismiss PII on the basis of the possibility of Black's world. However, Black's world is vulnerable to both the summing defense and the structure defense, and therefore no argument from Black's scenario to a denial of PII can be really deemed conclusive. My new counterexample solves this issue. If I am correct, in fact, all the lines of defense rehearsed against extant counterexamples to PII are unsuccessful against my scenario, which therefore puts unprecedented pressure on the Identity of Indiscernibles. I conclude this last Section by responding to a challenge according to which no branching world can contain two qualitatively indiscernible time-lines.

3.2 Other Counterexamples to PII

The Identity of Indiscernibles is a controversial principle, and has been challenged by numerous authors who believe it is possible that indiscernible entities do indeed exist. We have already seen some famous counterexamples to PII in Section 1.1.1, when we talked about examples of indiscernibles in Metaphysics and Ontology.

However, although Black's spheres, Ayer's sound tokens, and Strawson's chessboard are beyond any doubt among the most discussed counterexamples to PII, there are other counterexamples to the Identity of Indiscernibles which I didn't include in Chapter 1. This is because Chapter 1 was an introduction on indiscernibility, and as such it didn't engage with questions about the status of PII as a principle concerning the identity of individuals. I have therefore postponed the discussion of counterexamples such as Kant's droplets, Adams's twins, and Wüthrich's space-time points to the present Chapter, which focuses exclusively on whether the Identity of Indiscernibles is metaphysically necessary.

Before discussing the most common strategies that friends of PII usually employ to shield the principle against the threat of alleged counterexamples, I will then quickly go over these last challenges to PII.

3.2.1 Kant's Droplets

In his *Amphiboly of Concepts of Reflection*, an appendix of the Transcendental Analytic of his *Critique of Pure Reason*, Kant says:

> [...] in the case of two drops of water we can abstract altogether from all internal difference (of quality and quantity), and the mere fact that they have been intuited simultaneously in different spatial positions is sufficient justification for holding them to be numerically different.[1]

On the face of it, Kant is arguing against Leibniz's principle of the Identity of Indiscernibles by holding it possible for two drops of water to be qualitatively indiscernible. In Kant's example, in fact, the only difference between the two drops lies in their location, and therefore is, in Kant's absolute space, a non-qualitative difference.[2]

Hacking (1975, p. 251) interprets Kant's argument as an argument from abstraction. The argument starts from the premise that in the actual world one can easily find two drops of water. Then, the argument continues, one can easily abstract only these two drops of water from the actual world, ending up with a world where there are *only* two drops of water without internal differences. Since this world looks possible, then PII must fail.[3]

---

[1]This passage, found in Kant's *Critique of Pure Reason* (A263/B319), is quoted from Hacking (1975, p. 249).

[2]There is an incredibly rife literature surrounding Kant's *Amphiboly*. For a thorough discussion of Kant's main argument in the *Amphiboly*, see Ruffing et al. (2008). For a discussion about the correctness of Kant's account of Leibniz in the *Amphiboly*, see Bolton (2021), and Parkinson (1981). For more on Kant's account of identity and difference in the *Amphiboly*, see Brook & McRobert (1998).

[3]It is interesting to notice that this interpretation of Kant's counterexample is chal-

Hacking (1975, p. 251) takes issues with Kant's argument. For although he grants that Kant's abstraction is possible, he remarks that this doesn't suffice to reach Kant's conclusion. This is because still "[...] the question remains whether the result of this feat of abstraction is correctly described as having two indiscernibles in it. Simply to say so is to beg the question." (Hacking 1975, p. 251.)

3.2.2 Adams' Almost Indiscernible Twins

Another counterexample to PII comes from Adams (1979), who suggests that we can conclude that it is possible that indiscernible entities exist from the assumption that it is possible that *almost* indiscernible entities exist.

Adams takes this assumption to be innocuous. Virtually everyone, he remarks, would agree that it is possible that there is a world containing only two almost indiscernible spheres, one of which has a small chemical impurity the other lacks. To conclude from this that there is a world where these two spheres exist and are indiscernible, one would need a principle to the extent that the possibility of two entities existing in a given spatio-temporal relation is not affected by a slight change in their chemical composition.

However, Adams is well aware that such principle is controversial, and tries to advocate for a weaker one, according to which the possibility of two persons existing in a given spatio-temporal relation is not affected by a slight change in their mental events, provided these events have little if no effects

lenged by Nagel (1976, p. 46), who claims that Hacking's "[...] dialogue between [Kant] and [Leibniz] misrepresents the difference between the metaphysics of Kant and Leibniz. In extending our understanding in one respect, Hacking needlessly fosters misunderstanding in another".

on the reality surrounding the two persons.

Given this principle, Adams (1979, p. 17–19) argues as follows. It is uncontroversial that there is a possible world w where there are two almost indiscernible twins, call them Mike and Ike, the only difference between them being that on the night of their 27th birthday Mike was haunted in his dreams by a monster with ten horns, while Ike was haunted by a monster with only seven horns.

If this is possible, then it is also possible that both Mike and Ike didn't pay too much attention to their respective dreams, and therefore that having these dreams didn't cause any changes to Mike's and Ike's lives, nor to their physical surroundings.

Granted this, Adams argues, it would be foolish to believe that Mike couldn't have existed, or Ike for that matters, if the monster haunting him in his dream had had only seven horns. Therefore, we should conclude that there is a world w_1 where both Mike and Ike exist, and where they are indeed qualitatively indiscernible.

To discuss the details of Adams' argument against PII would take us too far afield, since the aim of this Section is only to introduce some other counterexamples to the Identity of Indiscernibles which were left out from Section 1.1.1. However, there are two aspect of Adams' counterexample which are worth pointing out.

The first is that the argument heavily depends on *transworld identity*, namely: identity between individuals across possible worlds. (See Adams 1979, p. 18.) Famously jilted by Lewis (1986) in favour of his counterpart theory, transworld identity has been defended, among others, by Kripke (1971) and Plantinga (1974; 2003). Adams' argument relies on transworld identity in order to show that in w_1, Mike and Ike are numerically distinct individuals. We can reach such conclusion, in fact, only if we hold it true that Mike and Ike in w_1 are identical, respectively, to Mike and Ike in w. Only if this identity statement is true we can then use the transitivity of identity or the necessity of identity to conclude that Mike and Ike are indeed two distinct individuals in w_1.[4]

The second is that Adams' argument depends on the further assumption that Mike could have been the twin dreaming of a seven horned monster while Ike could have been the one dreaming of a ten horned monster. This is because, Adams (1979, p. 18) suggests, two individuals differing only with respect to some modal properties could still be said to be qualitatively different. Therefore, for his argument to go through, Adams needs the further assumption that if the world w is possible, then also is the world w_2, where w_2 agrees with w in all respects except for the fact that in w_2 it is Ike who dreams of a monster with ten horns while Mike dreams of a monster with seven horns.

To support these two claims, Adams (1979, p. 18–19) argues that, if Mike and Ike exist as distinct individuals, this distinctess cannot depend on something that has not yet happened: for him, the identity of persons is always determined by their past and present, never by their future. If this is true, then when Mike and Ike were both 22 years old, they were qualitatively indiscernible in w too. (Which means that w is to some extent

[4]For more on the necessity of identity, see: Barcan (1947), Kripke (1971), and Wiggins (1965).

indeterministic.) Adams takes this to show that the numerical distinctness of Mike and Ike must be independent of the qualitative difference that arise eventually with their dreams.

3.2.3 Wüthrich's Space-Time Points

A final counterexample to PII comes from Wüthrich (2009), where it is argued that highly symmetric cosmological models in general relativity, when generally relativistic space-times are understood in a structural realist way, contain numerically distinct and yet indiscernible spatio-temporal points.

A little bit of stage setting. According to the structural realist interpretation of space-time, relativistic space-time consists of spatio-temporal points related by a set of suitable physical relations all deriving from the metric tensor associated to the relevant Lorentzian manifold. This kind of structural realism is called 'balanced' when it is further committed to the thesis that spatio-temporal points are as fundamental as the spatio-temporal relations they instantiate.[5]

Now, Wüthrich argues that balanced structural realists about generally relativistic space-times must hold that the only fact of the matter when it comes to the individuation of space-time points are the spatio-temporal relations they stand in to all the other points in the relevant Lorentzian manifold. As a consequence, they should endorse a structuralist version of

[5]Friends of this kind of balanced structural realism are Esfeld (2004), and Esfeld and Lam (2008). Pooley (2005) argues that a balanced understanding of structural realism is the most attractive view among all the versions of structural realism. For a thorough discussion of structural realism and its many strands, see Stachel (2005). A more controversial version of structural realism, according to which only spatio-temporal relations exist, is defended in French & Ladyman (2003).

PII according to which, for any space-time points x and y in a Lorentzian manifold, if x and y stand in all the same spatio-temporal relations to all the same space-time points in the manifold, then they are in point of fact identical. Wüthrich points out, however, that there are highly symmetric cosmological models which run against this version of PII, for they admit of distinct space-time points which are structurally indiscernible, in the sense that they stand in the same admissible physical relations to all the same space-time points in the relevant structure.[6]

If Wüthrich (2009, p. 1044–1046) is correct, then space-time points in some highly symmetric cosmological models are a counterexample to a structuralist version of PII according to which no two entities can share the same same automorphically invariant relational properties.

3.3 Defending PII

According to Hawley (2009), there are three strategies which can be employed to shield PII against putative counterexamples. These are: (1) the *identity defense*, which aims at finding an alternative description of the alleged counterexample to PII in which the two indiscernibles are identified as a single individual, (2) the *discerning defense*, which aims at finding some overlooked properties or relations that can help discern the alleged indiscernibles in the counterexample at hand, and (3) the *summing defense*, which aims at finding a consistent way of describing the scenario in the alleged counterexample as

[6] The details of why this is so are complex, and it is well beyond the purpose of this section to discuss them, even quickly. The interested reader is thus referred to Wüthrich (2009, p. 1044-1045).

a scenario containing only one simple scattered individual.

To the strategies discussed by Hawley I add a fourth one, which I call the *structure defense*. Employed by Hacking (1975) against Black's counterexample to PII, this last strategy aims at finding an alternative description of the scenario at hand in which a change in the definition of the background spatio-temporal structure allows to identify the two alleged indiscernibles.

In this Section, I discuss these lines of defense with the aim of showing that, in presenting a new counterexample to PII which is immune to all of them, I am not merely engaging in a stylistic exercise. For why even bother coming up with yet another counterexample, one may ask, when the majority of authors already believe PII is not necessarily true given that Black's two-spheres world is possible? The answer I am going to give is that, although it is true that Black's scenario is the go-to argument against PII, it turns out that Black's counterexample is vulnerable to both the summing defense and the structure defense. And therefore, any argument against PII from the possibility of Black's symmetrical world is far from water-proof. My counterexample, on the other hand, will be shown to be successful against *all* these lines of defense, therefore constituting a much more secure basis from which to argue against the necessity of PII.

3.3.1 The Identity Defense

Suppose you hold that PII is necessarily true, and suppose that some day a colleague of yours comes up with a scenario S containing two qualitative indiscernible individuals a and b, and tells you that since S seems indeed a possible scenario, then PII must not be a necessary truth after all.

If you are not willing to abandon your beliefs in the necessity of PII without having first tried to save it in some way, you might consider going for the so-called *identity defense* of PII. If you do, then what you have to do to save PII is to show that it is possible to give a description of S which is compatible with every qualitative claim made by your colleague, and yet in which $a = b$.

Now, if according to your colleague's description of S there is some spatial distance between a and b, then, if you want to leave the space-time structure of S unchanged, you will most probably have to argue that one and the same individual can be wholly located in two different regions of space at the same time.

One famous way to to this is to resort to the so-called 'Bundle Theory of Substance', according to which individuals are nothing over and above bundles of co-instantiated universals. If the Bundle Theory is true, then there is nothing controversial in claiming that one and the same individual can be simultaneously wholly located in two non overlapping regions of space. After all, since universals are capable of multi-locations and individuals are just bundles of universals, it would appear that also individuals should be capable of multi-location. (See also Section 1.3.1.)

This is, for example, how Hawthorne (1995) argues against Black's two-spheres world. If Hawthorne is right and Black's spheres, being individuals, are just bundles of co-instantiated universals, then Black's world can be consistently re-described as a world with only one multi-located sphere. (This is what Hawthorne argues, at least.) And if this is so then Black's counterexample should not worry the friend of PII anymore.

Against Hawthorne, however, Hawley (2006, p. 107–108) argues that even with the Bundle Theory in the background, without a working account of property-exclusion the description of the identity defense is inconsistent. To explain. The proponent of the identity defense, in a case like Black's two-spheres world, must endorse the following: at any time, the property 'being a perfect sphere' is multiply instantiated (for it is instantiated by the same sphere twice, one for each of the two spatial locations it wholly occupies), while the property 'being a two-spherical object' is uniquely instantiated (for the only object there is is indeed a two-spherical object). But since there is only one object in the situation we are considering, then said object is both a perfect sphere and two-spherical — and this is a contradiction. (Remember that, according to the proponent of the identity defense, the two spheres are not proper parts of the two-spherical entity.)

If inconsistencies of this sort were only relative to properties of shape and location, it could well be argued that, as a consequence of multi-location, we could not expect our concepts of 'location' and 'shape' to be left unharmed. However, as Hawley (2009, p. 107–108)(2009:107-108) suggests, such inconsistencies plague many more properties than we should be willing to concede.

Hawley does not consider this as a full blown *reductio* against the identity defense, and I agree with her in this respect. However, unless some suitably natural theory of property-exclusion is presented, one must accept that the identity defense cannot properly shield PII from spatial dispersal counterexamples like Black's.

3.3.2 The Discerning Defense

A second strategy you might use to protect PII against your colleague's counterexample is the so-called *discerning defense*. In case you decide to go with it, all you have to do is find some properties or relations that you can use to tell a and b apart within S — after all, it is not insane to think that your colleague might have overlook some properties or relations in their assessment of a and b's indiscernibility.

Recall again the second order formula expressing PII:

$$\forall x, y(\forall P(Px \leftrightarrow Py) \rightarrow x = y)$$

where x and y are individual variables, and P is a predicate variable ranging over properties of some specified kind. Given this regimentation of the principle, it easy to see that a scenario S is a counterexample to PII if and only if S satisfies the following formula:

$$\exists x, y((\forall P(Px \leftrightarrow Py)) \wedge (x \neq y))$$

which is a complex conjunction in the scope of an existential quantifier, and where again x and y are individual variables, and P is a predicate variable.

One way to think about the identity defense and the discerning defense is in terms of this formula: while the identity defense is an attempt to show that this formula is false in S by showing that its *second* conjunct is false, the discerning defense is an attempt to show that this formula is false in S by showing that its *first* conjunct is false in S.

The discerning defense of PII has been famously applied by Caulton and Butterfield (2012) against Black's two-spheres universe.[7] Caulton and Butterfield's (2012) argument rests on the notion of weak discernibility, discussed at length in Section 1.2.3.

As the reader will remember from Section 1.2, Quine (1976) identifies three degrees of discernibility: absolute discernibility, relative discernibility, and weak discernibility. We say that individuals x and y are absolutely discernible whenever there is a monadic property P such that $P(x)$ and it is not the case that $P(y)$. We say further that individuals x and y are relatively discernible whenever there is a relation R such that $R(x, y)$ and not $R(y, x)$. Finally, we say that individuals x and y are weakly discernible whenever there is a symmetric and irreflexive relation R such that $R(x, y)$ and not $R(x, x)$.

According to Caulton and Butterfield 2012, p. 50, Black's spheres are weakly discerned by the relation of 'being two diameters apart from —'. If this is true, then Black's spheres are discernible after all, and the friend of PII should not worry about Black's counterexample anymore. However, Lowe (2016, p. 53–57) convincingly argues that such relation (and in general *any relation* that would weakly discern Black's spheres) can only be used to show that there is some *non-qualitative property* by which we can tell the spheres apart. And this doesn't weaken Black's counterexample, for, the reader will recall, Black's scenario is supposed to work against a version of PII which is restricted to qualitative properties only. (This is also the version of PII that I will challenge in the next Section.) So the fact that we can distinguish Black's

[7] Other applications can be found in Saunders & Muller (2008) and Muller & Seevinck (2009) against alleged counterexamples to PII with indiscernibles elementary particles. For a recent challenge to such strategy, see Caulton (2013).

spheres by means of some non-qualitative properties should not bother us too much — and should also be expected, to some extent, if we claim that there are non-qualitative properties around.

In order to distinguish two entities by means of an irreflexive and symmetric relation, Lowe argues, we do not only have to presuppose their respective identities: we do actually need the identity facts we presuppose to show that the two relevant entities can be told apart. In a purely qualitative setting, Lowe argues, we can only say that (1) each sphere is two diameters apart from some sphere, and that (2) each sphere is at zero distance from some sphere. And this does not distinguish any of the spheres in Black's scenario.[8]

3.3.3 The Summing Defense

If your colleague has convinced you that neither the identity nor the discerning defense work against their scenario, you might still try to go with the so-called summing defense.

When dealing with alleged counterexamples to PII, the summing defense holds that the scenarios at hand can be consistently described as containing only one individual, which, unlike the one posited by the identity defense, is (1) scattered across space (and, if required, time), (2) mereologically simple, and (3) identical to the mereological sum of the alleged indiscernibles, had they existed and had a sum. (See Hawley 2009, p. 111–114.)

Scattered individuals are entities that somehow lack spatial unity. A

[8] An informal version of Lowe's criticism to Caulton and Butterfield can already be found in Black (1952). Although after some insistence from his imaginary interlocutor Black allows the possibility of naming his spheres, in fact, he warns that any attempt of using their names to discern them would be illegitimate within his scenario.

formal definition can be found in Chisholm (1984, p. 91), according to which a (material) individual x is scattered if and only if there are two (material) individuals y and z such that "x is composed of y and z and [...] no part of y is in direct spatial contact with any part of z."[9]

Although complex scattered individuals, like flock of birds, forests and cities, are most of the times innocuous entities, the requirement needed by the summing defense that the unique entity in the relevant scenario must be *mereologically simple* seems somehow to threaten the strategy's overall consistency. Consider again Chisholm's definition above, which clearly entails that, when x is a scattered individual, its scattered components are proper parts of it. This immediately turns any attempt to conceive of a scattered simple into an attempt to conceive an inconsistent object.

Undoubtedly, this alone does not constitute an argument against the summing defense, for Chisholm's definition may be challenged on independent grounds. Nonetheless, I find it really difficult to conceive of a scattered entity without conceiving any of its 'local', or 'connected' components as its *parts*. To do this, one would have to abandon, among others, the very attractive intuition that a scattered individual occupies distinct non-overlapping regions of space *in virtue of* having non-scattered parts which exactly occupy those regions.

Hawley seems to share some of these worries. In discussing the summing defense, she points out that "[t]he problem with scattered simples is that it

[9]Chisholm's (1984, p. 91) definition of 'composition' goes as follows: an individual x is composed of individuals y and z if and only if (1) both y and z are proper parts of x, (2) y and z share no proper or improper parts, and (3) all parts of x share some part with either y or z.

is hard to see what more could be required for the existence of an [individual] than the existence of a maximally connected portion of matter; that is, it is hard to see what prevents each of the spherical regions from exactly containing an [individual]." (Hawley 2009, p. 113) Hawley solves this problem by implementing PII in the equation: given, say, Black's two-spheres universe, the reason why the spherical region of space occupied by sphere x does not contain an individual (that is, the reason why x should not be considered as an individual) is that (1) there is another spherical region which is filled by matter (namely: the region occupied by sphere y that, like sphere x, should here not be considered as an individual in its own rights) together with the fact that (2) the Identity of Indiscernibles applies to the scenario at hand.

In the next Section, I will not challenge the summing defense on the basis of a particular definition of scattered individuals, nor on the basis of some very general intuitions about the relation between the notions of scatteredness and parthood. I will then assume, for the sake of the argument, that (1) Chisholm's definition above is indeed incorrect, and that (2) Hawley's way out is indeed a viable option for the friend of PII against Black's scenario. I will take scattered entities to be defined only by means of their location, since this definition does not beg the question against Hawley's summing defense. I take it in fact that the friend of the summing defense would accept that x is a scattered individual if and only if (1) x occupies two or more non-overlapping regions of space, and (2) x is not fully in any of the non-overlapping regions it occupies. Notice that with this definition of scattered individuals in play, Black's two-spheres universe, when not implemented with a specific account of 'parthood' and 'individuality', as well as a suitable theory of location, succumbs to the summing defense of PII.

3.3.4 The Structure Defense

Finally, in case not even the summing defense works and you are still not convinced by your colleague's extremely strong counterexample that PII is not necessarily true, you could try one last strategy: the so-called structure defense.

First proposed by Hacking (1975), this line of defense aims at finding a new description for S, in which some new spatio-temporal structure collapses the two original indiscernibles on one another. Given that the overall aim of the structure defense is to identify the two indiscernibles, it can be seen as a special case of identity defense. Given the way in which this identification is obtained, however, the structure defense is stronger than the usual identity defense (which argues for multi-located individuals), in that it doesn't display the kinds of inconsistencies which, as we have seen in Section 3.3.1, weaken this latter strategy.

Interestingly, Hacking (1975) doesn't think of the structure defense as a way to uphold PII in the face of counterexamples, for a scenario in which PII holds true is not preferable, *per se*, to a qualitatively equivalent scenario in which PII fails. Rather, Hacking uses the possibility of the structure defense as a way to show that for any spatio-temporal possible world w and any description d of w such that according to d, w is a counterexample to PII, one can find a qualitatively equivalent description d_1 of w such that, according to d_1, w is compatible with PII. Hacking's (1975) main point is that contemplating spatio-temporal worlds will never be sufficient either to establish or to refuse a principle like PII — which for him (1975, p. 255) is a "metaprinciple about possible descriptions". This is because, according to Hacking (1975, p. 255–256):

> Whatever God might create, we are clever enough to describe it in such a way that the Identity of Indiscernibles is preserved. This is a fact not about God but about description, space, time, and the laws that we ascribe to nature.

Although Hacking (1975) uses the structure defense against Ayer's scenario (see Section 1.1.1), it is easy to see how this defense could be used against Black's two-spheres world. Following Adams (1979, p. 15), one can give a qualitatively equivalent description of Black's scenario in line with the structure defense and hold that Black's two-sphere world is instead a world which contains only one sphere embedded in a non-Euclidean space which is so tightly curved that, by starting on the sphere and travelling on a straight line for the distance of two diameters, one would end up on the sphere again.

3.4 A New Counterexample to PII

In this Section I present a new counterexample to PII in the form of a world that admits of multiple incompatible time-lines. The scenario I set up, which I call the 'Disintegrating World', can be understood as a refinement of Black's (1952) famous two-spheres world. Unlike Black's scenario, however, the Disintegrating World can be shown to be successful against both Hawley's (2009) summing defense and Hacking's (1975) structure defense of PII.

In what follows, I first set the stage by quickly going over the features of branching worlds I need to set up my new scenario (Section 3.4.1). I then present the Disintegrating World (Section 3.4.2), and argue that it is a genuine possibility (Section 3.4.3). After discussing some issues about how identity and persistence can be understood to make sense in this world (Section 3.4.4), I argue in Section 3.4.5 that the Disintegrating World is successful against all the lines of defense of PII rehearsed in Section 3.3. This means that the Disintegrating World is a stronger counterexample to PII than Black's two-spheres world: as we have seen in Section 3.3, in fact, Black's scenario is vulnerable to both the summing defence (Section 3.3.3) and the structure defense of PII (Section 3.3.4). Finally, I discuss a possible challenge to the Disintegrating World coming from the intuition that any two incompatible time-lines must be qualitatively discernible.

3.4.1 Branching Worlds

In the famous episode 'Remedial Chaos Theory' of the series *Community*[10], a group of friends at a housewarming party is waiting for delivery. When the rider arrives, they throw a dice to decide who will go collect the food. This event creates six different time-lines, corresponding to its six possible outcomes. Each time-line is a possible continuation of the original time-line (where the party was taking place), and is spatio-temporally disconnected from all the other time-lines. This is an example of a 'branching world'.

More formally, a possible world w is a 'branching world' whenever w's time-line splits as a result of indeterministic events that have more than one

[10] Community, Season 3, Episode 4.

incompatible outcome. The number of distinct outcomes of an indeterministic event happening in a branching world determines the number of 'branches' that result from it. We say that an event e happening at time t in world w is *indeterministic* if (1) e has multiple incompatible outcomes, and (2) each of e's outcome has, at t, a non-null probability to happen. Loosely speaking, we say that events e_1 and e_2 are *incompatible* whenever it is impossible for them to occur together within a single time-line. In a branching world, time-lines diverge as a consequence of indeterministic events, and each time-line represent one of the multiple outcomes of the relevant event.

I suggest that some branching worlds can violate a version of the principle of the Identity of Indiscernibles according to which necessarily, no two entities can agree with respect to all their qualitative properties. Furthermore, I suggest that the counterexamples to PII we can set up using branching worlds fare better than the standard counterexamples to the principle, when tested again two common defenses of PII: Hawley's (2009) summing defense and Hacking's (1975) structure defense.

3.4.2 The Disintegrating World

In particular, consider a branching world $\mathcal{U}$ where at a certain time t an indeterministic event e happens. At any time t^- in the past of t, $\mathcal{U}$ closely resembles Black's (1952) two-spheres universe, apart from the fact that the spheres in $\mathcal{U}$ are 'continuous individuals', namely: individuals occupying continuous regions of space. (See Markosian 1998, p. 8.)

These spheres (call them Castor and Pollux) are at a distance of 10 miles from each other, and occupy two continuous non-overlapping regions of space r_1 and r_2. Following Cartwright (1975, p. 156–157), we say that a region of space r is continuous whenever it is not the union of two non-null separated regions r_i and r_j. At t, the original time-line of $\mathcal{U}$ branches into as many time-lines as the non empty subsets of the total region $r = r_1 + r_2$. In particular, for any subset r_i of r there is a time-line h_i such that, at any time t^+ in the future of t, the only occupied space in the entire universe according to h_i is r_i. Therefore, there are two time-lines h_1 and h_2 such that, at any time t^+, according to h_1 all the matter of the universe is located in r_1 while according to h_2 all the matter in the universe is located in r_2. I claim that the sphere in h_1 and the sphere in h_2 are distinct although qualitatively indiscernible individuals.[11]

A reasonable interpretation of what happens in $\mathcal{U}$ is that at time t some sort of indeterministic destructive event e takes place. This event could have either left Castor and Pollux untouched, or destroyed part of them in any possible way. From a purely spatial perspective, some of the (spatial) points occupied until t are suddenly emptied, and the conditions at t cannot determine which set of points is indeed going to be emptied. In general, among the many configurations we find in the future of t, exactly one is identical to the initial configuration. The others differ from it to varying degrees, from just slightly different ones (for instance, any configuration where only some spatial points have been emptied), to moderately and wildly different ones (such as configurations in which only few points have been left untouched).

[11] Clearly, h_1 and h_2 are not the only time-lines we can challenge PII with. Indeed, $\mathcal{U}$ contains at least infinitely many such time-lines.

3.4.3 A Genuine Possibility

I hold $\mathcal{U}$ is a genuine possibility. First, $\mathcal{U}$ is a consistent scenario, in the sense that its qualitative arrangement does not entail the existence of any contradictory entity. Second, the overall scenario is clearly conceivable, in the sense of Yablo's 'philosophical conceivability', i.e. the kind of conceivability that "involves the appearance of possibility" (Yablo 1993, p. 7). Finally I hold that, if you believe that Black's (1952) two-spheres universe is a genuine possibility, then you should also believe that $\mathcal{U}$ is a genuine possibility.[12]

Here is why. When we consider $\mathcal{U}$ as a sum of instantaneous configurations, we can see that any of them is undeniably so similar to Black's scenario that, were one to consider Black's a genuine possibility, there would be no reasons to hold that $\mathcal{U}$ is not. The only relevant difference between each of $\mathcal{U}$'s single instantaneous configurations and Black's scenario is in fact the location of the relevant matter. However, given the nature of the individuals involved (i.e. aggregates of matter, much more similar to mathematical solids than to ordinary individuals), facts about location can hardly yield a difference with respect to the distinct configurations' possibility. I conclude that $\mathcal{U}$ is at least 'locally' possible.

When we consider $\mathcal{U}$'s time-lines, we see that the only thing that is happening within any of them is that some previously occupied spatial regions becomes empty after t.[13] A way to describe this is perhaps to say that, gen-

[12] Among others, Adams (1979), Lowe (2003) and Rodriguez-Pereyra (2017) all agree that Black's world is genuinely possible.

[13] The limiting case being the only time-line where no region is emptied as a result of e. This case is, however, uninteresting, and the argument I present is independent from the assumption of its existence.

erally speaking, $\mathcal{U}$'s time-lines witness the annihilation of some matter. And while it is true that any world in which mass is not constant is not nomologically possible, we are only challenging an interpretation of PII according to which the Identity of Indiscernible is unrestrictedly valid. This means that we only require $\mathcal{U}$ to be metaphysically possible.[14] And thought-experiments involving the annihilation of matter aren't new. A famous example involves Descartes and his left leg, and constitutes the main argument van Inwagen (1981) presents against the so-called 'Doctrine of Arbitrary Undetached Parts'. Other examples can be found in the literature on instantaneous temporal parts. (See Effingham 2012). Moreover, it is a well known facts that in Newtonian worlds, matter can appear and disappear instantaneously. I conclude that any of $\mathcal{U}$'s time-lines, considered singularly, is a genuine possibility.

Taken together, these considerations suggest that $\mathcal{U}$ is a genuinely possible world, provided of course (1) we have some reasons to believe that branching worlds are possible, and (2) we have some reasons to believe that genuinely indeterministic events are possible. I hold that there is no philosophically interesting reason to deny (1) since, as it is often assumed, branching worlds are genuinely possible. (See Belnap 1992, Belnap & Green 1994, and Barnes & Cameron 2011.) Finally, I take the fact that according to Quantum Mechanics the actual world is indeterministic to be enough to justify (2).

[14]The notion of metaphysical possibility I hold to be meaningful has been defended among others by David Lewis, Saul Kripke, Alvin Plantinga and Jonathan Lowe, and it has recently been challenged by Ladyman and Ross (2007).

3.4.4 Identity & Persistence

Still, someone may worry that branching worlds are incompatible with our intuitions about persistence, resulting in inconsistent descriptions of reality when everyday situations involving indeterministic events are considered. For instance, consider a scientist who wants to measure the spin of an electron e. According to Quantum Mechanics, measuring the spin of an electron is an indeterministic event with two possible outcomes: as a result of the measurement, either the electron is in the spin-up state or it is in the spin-down state. These outcomes are incompatible, and they have both a positive probability to occur.

If we are in a branching world then the measurement of the electron's spin creates two distinct time-lines h_1 and h_2. According to h_1, after the measurement the electron e is in the spin-up state; according to h_2, e is in the spin-down state. This entire story, however, seems to involve only *one* electron: the electron in the spin-up state, as well as the electron in the spin-down state, are the *same* electron that has been measured by the scientist. At first sight, this might strike as a violation of the transitivity of identity. The electron in h_1 must be clearly distinct, one might hold, from the electron in h_2, for only the first has the property 'being in spin-up state'. But if the electron in h_1 is the same as e, and if e is the same as the electron in h_2, then the electron in h_1 and the electron in h_2 must be one and the same electron, since identity is transitive. In other words, the example is one in which, after a certain interval of time δt, the electron e has spin-up in h_1 and spin-down in h_2. Since δt is a well-defined interval, e has both spin-up and spin-down at the same time. The conclusion is that, no matter what, the electron in h_1 and the electron in h_2 cannot be both identical to e.

According to this argument, only one of the electrons in h_1 and h_2 has the property 'being in spin-up state', which means that they are qualitatively distinct. And since no two qualitatively distinct individuals can be numerically identical, the argument concludes that the two electrons must be distinct. We can resist such conclusion by holding that, within the relevant scenario, no electron has the property 'being in spin-up state' *simpliciter*. Given the indeterministic nature of the measurement, in fact, all the properties relative to the electron's measured spin must be relativised to the relevant time-lines. There is no 'being in spin-up state' property. Instead, the only property the electron e has with respect to the spin-up eigenvalue is 'being in spin-up state relative to h_1'. We can say the same for the property 'being in spin-down state'. In the example above, the electron in h_1 has the property 'being in spin-up state relative to h_1', and the electron in h_2 has the property 'being in spin-down state relative to h_2'. These two properties are no longer mutually exclusive. Furthermore, the electron in h_1 has the property 'being in spin-down state relative to h_2', and the electron in h_2 has the property 'being in spin-up state relative to h_1'. There is no property that the first electron has and the second lacks. Therefore, it is clearly possible for the electron in h_1 to be the same as e, and for e to be the same as the electron in h_2, and the transitivity of identity is preserved.[15] Furthermore, we should note that the use of a single time interval to argue that the electron in h_1 and the

[15] Similarly, one cannot argue that the electrons must be distinct for they are in different time-lines. The electron in h_1 has in fact the property 'being in h_2 relative to h_2', and the same holds for the electron in h_2 relative to h_1.

electron in h_2 have incompatible properties *at the same time* is problematic in the context of branching worlds, for time itself, in this setting, is relative to time-lines.

Also, current accounts of persistence can make perfect sense of our electron's identity over time. Here I just consider three-dimensionalism and four-dimensionalism, for clearly a sequentialist account of persistence would have no problem in explaining e's diachronic identity. [16]

Three-dimensionalism is the thesis that material individuals do not have temporal parts, and persist through time by existing at different times and being wholly present at any time they exist. Three-dimensionalists can explain the identity facts regarding e by saying that, time being relative to time-lines, e has nothing more than different properties at different times relative to different time-lines. As they hold that one and the same individual can have incompatible properties at different times without contradiction, they should for the present purposes hold that: (1) the electron e has the property 'being in an indeterminate spin state' at any time t before the measurement relative to the original time-line, (2) e has the property 'being in spin-up state' at some later time t^1 after the measurement relative to h_1, and (3) e has the property 'being in spin-down state' at some later time t^2 after the measurement relative to h_2.

[16] See: Geach (1965, 1967) and Van Inwagen (1990) on three-dimensionalism; Lewis (1986), Armstrong (1980), and Noonan (1980) on four-dimensionalism; and Chisholm (1976), Lewis (1968) and Varzi (2003) on sequentialism. Orilia (2006, p. 206) characterises sequentialism as the thesis that "[...] there are temporal parts ordered by temporal relations like before and after, but there is no *objective* relation of genidentity that relates some of these parts in such a way that they come to constitute one ordinary [individual] that perdures in time. We can however speak of some sequences of these temporal parts *as if* they were [individuals] perduring in time, given the choice of a reidentification criterion".

Notice that this does not entail a contradiction, even when t^1 and t^2 are simultaneous in the sense that the length of the time interval δ_1 separating t from t_1 is the same as the length of the time interval δ_2 separating t from t_2. Even if t_1 and t_2 are simultaneous, in fact, they are distinct moments in time, for one is a point in h_1 and the other a point in h_2.[17]

On the other hand, four-dimensionalism holds that material individuals persist through time in virtue of having different temporal parts, and that one and the same individual passes from being P to being P^* (where P and P^* can be incompatible properties) in virtue of having some temporal parts which are P and some (successive) temporal parts which are P^*. Four-dimensionalists can make sense of our example by holding that, within a branching universe, the same individuals can have temporal parts in multiple time-lines. Although this might not seem intuitive, I believe it is actually a natural extension of the classical four-dimensionalist account. In a universe without incompatible time-lines any individual has a unique trajectory in space-time, and the history of that individual is completely specified by its unique trajectory. In a branching universe, the trajectory of a given individual branches as space-time branches, and a unique trajectory can specify at

[17] The friend of three-dimensionalism has two options here: either they hold that properties are by their very nature relative to the relevant spatio-temporal framework (in which case it is not a contradiction for an individual to instantiate incompatible properties in distinct time-lines), or they hold, expanding on an already well-known three-dimensionalist position, that what we usually think about as properties are indeed relations. Here, the property 'being in spin-up state' could be thought as a three-place relation holding between an individual, some moment in time, and some time-line. Alternatively, a more conservative solution is available. Since, according to Belnap, time-lines are sets of spatio-temporal points, there is room for a view that moments in time are themselves already distinguished by the membership relation between times and time-lines. Under this account, properties like 'being in a spin-up state' could be interpreted as binary relations, holding between individuals and moments in time.

most one of the multiple histories of the individual in question. By holding that an individual can have temporal parts in all the time-lines of a branching universe, the four-dimensionalist is doing nothing more than to adapt their theory to the topological properties of the universe the individuals they are talking about inhabit. And well they might: for the location of individuals in space and time, and therefore their history, necessarily depends on the topological features of the universe they are in. At this point it should be clear that, whenever there are multiple time-lines and some individual whose trajectory encompasses them all, its temporal parts must be relative to time-lines. Said individual will have temporal part s_1 at time t^1 in h_1, temporal part s_2 at time t^2 in h_2, and so on. In particular, e will have a temporal part e_1 which is in spin-up state, and a temporal part e_2 in spin-down state. The indeterministic nature of the measurement of e's spin will be then given by the fact that e_1 and e_2 are in distinct time-lines. Under such account, an individual will be nothing more than the sum of all of its temporal parts, as in the classic four-dimensionalism without branching. The fact that in a branching universe an individual's temporal parts might inhabit distinct time-lines is not contradictory, for all their properties are accordingly relativised.

Another way for the four-dimensionalist to make sense of our electron's identity over time is by adopting an account of fission in line with Lewis (1976). Accordingly, they can claim that our example involves *two* four-dimensional electrons which share an initial temporal segment.[18]

[18] For an account of fission within three-dimensionalism, see Merricks (1997).

3.4.5 Meeting PII's Defenses

As I have mentioned in Section 3.3, many of the arguments one finds in the contemporary literature against PII all take off from the assumption that Black's world is indeed genuinely possible. From this, it follows that the strength of these arguments is a function of the overall strength of Black's counterexample. In other words: the more conspicuous is the number of lines of defense of PII which are successful against Black's scenario, the weaker any argument against PII which depends on its possibility will be.

Now: while successful against the identity and the discerning defense, Black's two-spheres world is vulnerable to both the summing and the structure defense of PII. And while one could independently argue against these two last lines of defense, it is obvious that if there were a counterexample to PII which was successful also against the summing and the structure defense, the best strategy to argue against PII would be to start from *that* counterexample instead — for this would make for a stronger argument than the one whose premises are (1) that Black's world is possible, (2) that the identity defense is mistaken, and (2) that so is also the structure defense.

This is what the Disintegrating World promises to do, and this why I think it is worthy of consideration. In what follows I argue that, unlike Black's scenario, the Disintegrating world is successful against *all* the lines of defense of PII discussed so far.

Meeting the Identity Defense

As we saw in Section 3.3.1, the identity defense consists in finding a consistent description of the alleged counterexample to PII as a situation involving only *one* entity. Again, this is usually taken to entail that individuals are capable of multi-location, and for this reason the identity strategy is usually held together with the Bundle Theory of Substances, which derives the possible multi-location of individuals from the almost uncontroversial possible multi-location of universals.

Since $\mathcal{U}$ contains (infinitely) many distinct instantaneous configurations involving pairwise indiscernibles, we can dismiss the identity defense if we can show that there is at least one such configuration that cannot be consistently described according to the standards of the strategy. (The same also applies to the other defences of PII.)

Consider, then, $\mathcal{U}$'s qualitative arrangement at any time before t. According to the identity defense, there is only one multi-located sphere in $\mathcal{U}$, fully occupying the two non-overlapping spatial regions r_1 and r_2. However, as per Hawley (2009, p. 107–108), since the friend of the identity defense must endorse that at any time in the past of t the property 'being a perfect sphere' is multiply instantiated while the property 'being a two-spherical object' is uniquely instantiated, they are committed to the existence of an object which is both a perfect sphere and two-spherical. And this is a contradiction.

As we have already seen, this is not the only contradiction the identity defense must endorse. Say, for example, that any of the spheres in $\mathcal{U}$ has mass m. Then, again, the friend of the identity defense must endorse that the property 'having a mass m' is multiply instantiated while the property 'having a mass which is the double of m' is uniquely instantiated in $\mathcal{U}$ before t. But since there is only one multi-located individual, then this individual must have mass m and $2m$ simultaneously.

I agree with Hawley (2009, p. 108) that it is possible to endorse a suitable account of property exclusion that would allow the friend of the identity defense to avoid these contradictions and present a good case against the possibility of the Disintegrating Universe. However, any such account would also make the identity defense successful against Black's scenario. Therefore, I conclude that if Black's scenario is successful against the identity defense, also the Disintegrating Universe is.

Meeting the Discerning Defense

Unlike the identity defense, the discerning defense aims at finding some properties or relations which can be used to discern between the alleged indiscernible entities in the relevant scenario. So the question is: can we find, for any pair of allegedly indiscernible individuals in $\mathcal{U}$, some property or relation which can tell them apart?

If we can, then it should be clear that $\mathcal{U}$ is not a real threat to PII after all. However, if we can show that there is even only one pair of entities in $\mathcal{U}$ which are not even weakly discernible, then we can dismiss the discerning defense as unsatisfactory once and for all. And as with the identity defense, I hold that to this end, one needs just to look at Castor and Pollux: the two spheres in $\mathcal{U}$ at any time before t.

Clearly, in fact, Castor and Pollux are not absolutely discernible, at least with respect to qualitative properties: they have the same geometry, are composed of the same kind of matter, and their internal structure looks exactly the same. Arguably, they have also the same mass and weigh the same. Furthermore, they have the same number of conceptual parts, and there is no good reason to think they have a different number of real parts. (Clearly, they have different conceptual part, as well as different real parts — that is, if they have real parts at all. However, such differences are not qualitative, since they depend on the identity conditions of the entities we take to be parts of the spheres.)

Similarly, they are not relatively discernible, since there is no asymmetric relation holding between Castor and Pollux. All the spatio-temporal relations obtaining between them are symmetric, as are all the relations whose definition employs talk of physical magnitudes or mereological concepts.

The only possibility left to the friend of the discerning defense is therefore weak discernibility. However, weak discernibility won't work either, and here is why. We have seen in Section 3.2.2 that the discerning defense is unsuccessful against Black's (1952) two-spheres universe, for any relation that would weakly discern Black's spheres can only be used to show that there is some non-qualitative property by which we can tell the spheres apart. (See Lowe 2016.)

The same hold with our spheres, Castor and Pollux. And this is not just because the initial segment of $\mathcal{U}$ up to t is indistinguishable from an initial segment of Black's two-spheres world of the same duration. More generally, it is because Lowe's (2016) argument against weak discernibility as a mean of *qualitative* difference applies to any scenario with indiscernibles. Therefore, if

Lowe is right, the discerning defense is not strong enough to defend PII from the Disintegrating World. That is: with respect to the discerning defense, Black's world and the Disintegrating World stand or fall together.

Although in arguing against the identity and the discerning defense we have only considered the qualitative arrangement of $\mathcal{U}$ in the past of t, it is easy to see that the difficulties that the two strategies run into would only be worsened by considering any suitable qualitative arrangement (or pair thereof) in the future of t.

Meeting the Summing Defense

Recall that, according to Hawley (2009, p. 111–114), the summing defence consists in finding a suitable alternative description of the alleged counterexample to PII as containing only one mereologically simple, scattered object, identical to the mereological sum of the alleged indiscernibles, had they existed and had a sum.

Recall further that, in order not to beg the question against the summing defense, in Section 3.3.3 we defined scattered objects as those which occupy two or more non-overlapping regions of space, without at the same time being wholly in any of the non-overlapping regions they occupy.

Finally, notice that $\mathcal{U}$'s initial configuration is not sufficient to challenge the summing defence. Were our universe to implode before t, in fact, the friend of the summing defence could describe it as containing only one simple object scattered across two spherical regions of space — and at that point we would not be able to challenge Hawley's defence without begging the question against the whole strategy. For the same reason, Black's two-spheres universe fails against the summing defence of PII.

With this in mind, consider the instantaneous configuration including h_1 and h_2 at any time in the future of t. (Remember: h_1 is the time-line where only Castor survives, while h_2 is the time-line where only Pollux survives.) According to the summing defense, such instantaneous configuration contains only one scattered object, call it Callux. Clearly, was Callux a composite object, we could still challenge PII by referring to its indiscernible, and yet distinct, parts. In such case, in fact, Callux would at least have some proper parts, and for any part x Callux might have in h_1 there would be a corresponding entity y in h_2 so similar to x that it would be impossible to give a non *ad hoc* reason why one should consider x, and not y, a part of Callux. In case an argument was made by the friend of PII that none of the parts of Callux was fully inside one of the time-lines, we could then ask, about any of its parts, if it is a scattered individual or a multi-located one, thus running the same argument again. As a consequence, Callux must be *simple*.

Now, while we didn't have any reasons to deny *a priori* the possibility of there being mereologically simple individuals scattered across space and time, this is not so with mereologically simple individuals scattered across distinct space-times. Say that an individual x is scattered$_1$ whenever it is scattered across space and time, while it is scattered$_2$ when it is scattered across different space-times. From a methodological point of view, to include scattered$_2$ entities in one's ontology is not an unsubstantial move. The possibility of scattered$_2$ entities does not follow from the possibility of scattered$_1$ ones. Furthermore, *simple* scattered$_2$ entities seem to run against our intuitions about material individuals. Notice in fact that when applied to Castor and Pollux in the time interval $(t, ...)$, the summing defense entails that both h_1 and h_2, taken in isolation, are empty time-lines: and this doesn't seem right. That no material individuals can be found neither in h_1 nor in h_2 does not only contradict common intuitions about what material individuals are, and the identity and individuation conditions that govern their existence. More importantly, it contradicts the fact that when there is some spatial region occupied by matter, and every other spatial region is empty, we can identify a material individual. Now, the friend of PII has definitely lost my understanding of the content of their argument. I doubt that by 'material individual' we are meaning the same thing anymore. And if we are not, I cannot understand what a material individual is supposed to be, under such account.

We cannot yet dismiss the summing defense as unsatisfactory, for the friend of PII could still argue that, if our branching world is meant to show that the principle is invalid, challenging it by considering only an instantaneous temporal part of $\mathcal{U}$ is insufficient. However, considering the entire history of $\mathcal{U}$ with respect to h_1 and h_2 will not get the summing defense out of trouble. Here is why. When considering our branching world, we see that it presents two different challenges to PII.

The first is in the unique time-line in the past of t, and the second is in its future, when considering suitable sets of incompatible time-lines. While, as we said, the summing defense easily applies to the first configuration taken in isolation, the same cannot be said for the second. What this means is that, if it is true that the only way out of the impasse is, as the friend of PII at this point holds, to consider the scenario in its entirety, then the summing defense is bound to give a suitable non contradictory description of the entire universe — and part of this description must be that such universe contains, in the past of t, nothing more than a single mereologically simple individual, scattered across two spherical regions of space.

Now, the possibility we had to explain $\mathcal{U}$'s history by holding that some of the matter in it gets destroyed at t cannot be available to the friend of PII, since they cannot maintain that some parts of their initial simple individual have been annihilated as a consequence of the indeterministic event at t. Furthermore, they cannot explain the situation by saying that the only change resulting from t is that some simple individual has passed from being scattered across space and time, to being scattered across multiple (disconnected) space-times, for that would run into contradiction with the intuition that h_1 and h_2 are not empty time-lines.

At this point the friend of PII is left with only two alternatives, and both run against PII. They might point out that what we called Castor and Pollux are just alternative outcomes of a process of (instantaneous) contraction undergone by an simple scattered individual s at t, to which we can respond that, even when considered as two incompatible states of s, Castor and Pollux are still clearly distinct and indiscernible. The only possible explanation left is one according to which no individual in $\mathcal{U}$ survives t. According to this interpretation, s disappears at t and some other individuals appear right after it. Although this solution avoids the friend of PII the problem of the alleged parts of s, it does not yet succeed in saving the Identity of Indiscernibles. What appears after t cannot in fact be just a unique simple scattered individual.

Meeting the Structure Defense

Finally, let's consider the structure defense. As we have seen in Section 3.3.4, the structure defense consists in identifying the alleged indiscernibles by identifying the spatio-temporal regions they occupy, thereby changing the geometry of the underlying spatio-temporal framework.

Is this defense successful against the Disintegrating World? If we just consider what happens before t, then yes: the friend of PII has all the rights to say that, at any time before t, $\mathcal{U}$ contains only one sphere, ten miles apart from itself, in a curved space-time framework. Notice that we cannot just reply that the spheres are two, for then we would just be begging the question against PII. And we cannot even reply that a change in spatio-temporal framework must yield a change in possible worlds: for (1) in order to challenge PII we need our descriptions to remain on a purely qualitative level, and by definition any purely qualitative description of a given scenario will underdetermine the spatio-temporal framework of the world in question, and (2) we really don't want to commit to the absurdity that for any given world w, there are infinitely many distinct worlds which agree with w about everything else including the *relative* positions of their entities, except for the *absolute* positions of their entities. This is why Black's scenario can indeed be alternatively described in a way that is compatible with PII, failing against the discerning defense.

Let's then consider what happens in the future of t. Here, we find as many distinct time-lines as the subsets of $r_1 + r_2$. In particular, for any pair of points p_1 and p_2 in r_1 and r_2 respectively, there are two time-lines h_i and h_j such that: in h_i the only occupied point is p_1, and in h_j the only occupied point is p_2. Since the non-extended objects in h_i and h_j are a counterexample to PII and since this holds for any pair of points in r_1 and r_2, the only way to successfully apply the structure defense (thereby avoiding any possible counterexample to PII in the future of t) is to identify all points in $r_1 + r_2$ with a unique point p. However, the resulting description wouldn't be a good description of $\mathcal{U}$, for it wouldn't agree with our original description in one qualitative key aspect: that the individuals in $\mathcal{U}$ are extended. Since the structure defense aims to find a new description of the old scenario that agrees with the previous description in all qualitative aspects while entailing new quantitative aspects, this defense cannot be successfully applied to $\mathcal{U}$.

3.4.6 Indeterminism in the Disintegrating World

In this last Section I consider a methodological objection to the construction of the Disintegrating World, stemming from the thesis that no two time-lines h_i and h_j can diverge that are not, in some sense, qualitatively different. This thesis, which I call the 'Qualitative Collapse Principle' (QC for short), is meant to generalise the intuition that, were we asked to consider two time-lines h_i and h_j such that (1) all the objects in h_i uniformly move in the same direction at a given speed c, and (2) all the objects in h_j uniformly move in the same direction at speed d different from c, then we wouldn't actually be considering *two* time-lines. In other words, imagine God had a giant button, by pressing which He would set the entire universe in linear motion. Imagine further that it was indeterminate at the time God pressed the button whether the universe would be set as to uniformly travel towards direction $\vec{v}$ at uniform linear speed c or at uniform linear speed d greater than c. Even if the divine act of pressing the giant button had at least two distinct incompatible outcomes, according to QC this fact alone wouldn't be enough for distinct incompatible time-lines to be generated — which in turn means that pressing the button was not an indeterministic event after all.[19]

I hold that QC is wrong, for it relies on an absolute notion of indeterminism. To see why, suppose that, under ideal conditions, the tossing of a fair coin is an indeterministic event. Suppose you have a fair coin, with two *heads* instead of the usual *heads* and *tails*. When you toss the coin, there is of course no doubt that the outcome will be *heads*. However, this does not mean that tossing the unusual coin is a deterministic event. For since a coin does have two sides, and it does so necessarily, there are *two* ways for the coin to show *heads*, not just one. And if it is assumed that the number of outcomes of an indeterministic event happening in a branching world determines the number of branches that result from it, then there must be two branches resulting from tossing the coin, as well as there must be infinitely many branches resulting from God pressing the giant button. A similar case can be made for the Disintegrating Universe: for since r is extended, all its distinct sub-regions exist, and they do so necessarily. And this is why the time-lines h_1 and h_2 don't collapse: there were many possibilities for the outcome of e to be, and among them there were (1) the possibility of r_1 being

[19] Note that the fact that the event in question turns out not to be indeterministic is in no way dependent on God's omniscience or omnipotence. Where I have imagined God, one can clearly imagine a mad scientist.

occupied, and (2) the possibility of r_2 to be so.

Chapter 4

PII & Haecceitism

4.1 Introduction

In this Chapter I argue that if a popular version of the Identity of Indiscernibles relativised to ordinary spatio-temporal entities is not necessarily true, then Haecceitism follows — where Haecceitism is the view that there are maximal possibilities which include all the same qualitative possibilities, and yet differ with respect to the non-qualitative possibilities they include. This goes against the common intuition that Anti-Haecceitism is compatible with this version of the Identity of Indiscernibles, which I call PII-WB, being only contingently true.

The argument I set up in this Chapter is interesting in many respects. First, it shows that in *any* modal framework there is a connection between the number of worldbound ordinary spatio-temporal individuals, and the number of overall possibilities. Second, it has repercussions for the tenability of some philosophical theories, like Generalism, which are usually interpreted as entailing Anti-Haecceitism whilst at the same time being compatible with the claim that PII-WB is not necessarily true. If I am correct, Generalism and similar philosophical accounts turn out to be inconsistent.

Finally, my argument lends very strong support to at least some weak form of Haecceitism, given that the majority of authors today find counterexamples to the Identity of Indiscernibles extremely convincing, and many philosophical positions have been and continue being criticised on the basis of their commitment to PII-WB.

This Chapter is structured as follows: I introduce Haecceitism and the Identity of Indiscernibles in Section 4.2 and Section 4.3 respectively. Then, drawing on a result from the philosophy of Quantum Mechanics, which I survey in Section 4.4, I present my main argument in Section 4.5. I discuss my conclusions and their relevance for the current philosophical debates on Haecceitism and PII in Section 4.6.

4.2 Haecceitism

If you believe that the world could have been exactly as it is, except for the fact that I could have had all the qualitative properties you actually have and you could have had all the qualitative properties I actually have, then you are a Haecceitist.

Haecceitism holds that there are ways the world could have been that differ from the way the world actually is only with respect to some non-qualitative properties or facts.[1]

[1] For discussion, see: Adams (1979), Cowling (2017, 2023), and Lewis (1986).

Non-qualitative properties are properties which depend, in one way or other, on the identity of some specific individual. 'Being Hillary Clinton' or 'standing next to Joe Biden' are common examples of non-qualitative properties. Qualitative properties, on the other hand, do not depend on any individual: 'being someone's employee', 'being extended' and 'being in love with someone' are all examples of qualitative properties.[2] Similarly, we take the fact [that someone is tall] as qualitative, and the fact [that Boris Johnson is tall] as non-qualitative.

4.2.1 Haecceitism & Possibilities

The same distinction applies to possibilities, which are commonly understood as ways the world could be/could have been. We say that a possibility is qualitative when it does not depend on any specific individual, non-qualitative otherwise. Accordingly, the possibility *that aliens exist* is qualitative, while the possibility *that Donald Trump was a song-writer* is non-qualitative, since it depends on one specific individual: Donald Trump.

Usually, we distinguish between *maximal* and *non-maximal* possibilities: maximal possibilities are *total ways* the world could have been, while non-maximal possibilities are *less than total ways* the world could have been. (For elaboration, see Stalnaker 1984.) For instance, the possibility *that Joe Biden lost the election* is non-maximal, for it tells us nothing about the world

[2] The jury is still out on how to define qualitative and non-qualitative properties. Rodriguez-Pereyra (2006), for instance, holds that non-qualitative properties are those depending on the *identity* of some specific individuals, while according to Hawley (2009), properties are non-qualitative whenever they depend on the *existence* of some specific individuals. For the present purposes, an intuitive understanding of qualitative and non-qualitative properties will suffice. I will propose an account of qualitative and non-qualitative properties in Chapter 5.

apart from what happened to Joe Biden, and perhaps some other facts which follow from his electoral loss.

Finally, we say that some possibilities *include* other possibilities, and that a maximal possibility includes both qualitative and non-qualitative possibilities. An example: the possibility represented by the actual world is by definition maximal, and includes both qualitative possibilities, like the possibility *that atoms are composed of protons*, and non-qualitative ones, like the possibility *that Hillary Clinton lost against Donald Trump*.

Cowling (2017, p. 4172) notes that there are different ways to understand the relation of inclusion between possibilities. Those who believe that possibilities are propositions are likely to understand inclusion as an instance of entailment, while those who think of possibilities as sets of propositions usually understand inclusion in a set-theoretical way. Here, I will remain neutral on how we should understand the relation of inclusion, for nothing I'm going to say hinges on one particular interpretation of this relation between possibilities.

With this in mind and following Cowling (2017, p. 4172), we define Haecceitism as follows:

> **Haecceitism:** There are maximal possibilities which include all the same qualitative possibilities, and yet differ with respect to the non-qualitative possibilities they include.

In what follows, I will argue that Haecceitism so defined follows from the thesis that a popular version of the Identity of Indiscernibles, according to which no two distinct ordinary spatio-temporal entities can agree with respect to all their qualitative properties, is not necessarily true.

4.2.2 Haecceitism & Possible Worlds

Some authors identify possibilities with possible worlds, while others hold that one and the same possible world can represent distinct possibilities.[3] The argument I will set up will not require any decision on this matter: whether possibilities and possible worlds are one and the same or not will not significantly influence any of the argumentative steps I will present in Section 4.4.2. Therefore, I will here refrain from taking a side in the debate.

It is interesting to notice, however, that those who hold that possibilities are not the same as possible worlds, and that one and the same possible world can represent distinct possibilities, can endorse Haecceitism as we have defined it in Section 4.2.1 without having to endorse the thesis that there are distinct possible worlds which represent distinct maximal possibilities which in turn include the same qualitative possibilities, and differ only with respect to the non-qualitative possibilities they include.

This second thesis, which Cowling (2017, p. 4174) calls *Possible Worlds Hecceitism*, is quite independent from *Possibilities Haecceitism* (which is Cowling's name for the thesis we have defined in Section 4.2.1). Unlike Possibilities Haecceitism, in fact, Possible Worlds Haecceitism is a thesis about the metaphysical relation obtaining between possibilities and possible worlds, entailing a one-to-one correspondence between them.

It follows that while Possible Worlds Haecceitism entails Possibilities Haecceitism, the converse doesn't hold. And this means that it is still con-

[3] See, among others, Lewis (1986, pp. 230–231).

sistent to hold the latter thesis whilst rejecting the former. As we will see in Section 4.3.3, Lewis (1986) famously endorses Possibilities Haecceitism and rejects Possible Worlds Haecceitism.

In what follows I will only consider Possibilities Haecceitism, and refer to it as Haecceitism *simpliciter*.

4.2.3 Haecceitism & Haecceities

Cowling (2023) notices that Haecceitism can be taken to entail the existence of *haecceities*. Sometimes also called 'thisnesses' or 'individual essences', haecceities are identity properties of the form '$\lambda x[x = c]$', for 'c' some individual constant. The property 'being Joe Biden' is a haecceity, and so is the property 'being Napoleon'. The property 'being self-identical', on the other hand, is not a haecceity. With Adams (1981, p. 4), we can characterise a haecceity as:

> the property of being a particular individual, or of being identical with that individual. It is not the property we all share, of being identical with some particular individual or other. But my [haecceity] is the property of being me; that is, of being identical with me. Your [haecceity] is the property of being you. Jimmy Carter's [haecceity] is the property of being identical with Jimmy Carter (not: of being called "Jimmy Carter"); and so forth.

Often dismissed as 'creatures of darkness' due to their alleged unintelligibility, haecceities are still useful to explain differences in non-qualitative properties and relations.[4] They are also sometimes employed as grounding bases for the individuation of individuals and universals, in contexts where individual-identity and universal-identity are thought of as primitive.

Haecceitists can resort to haecceities to explain the distinctness of possibilities which disagree only with respect to the non-qualitative possibilities they include. Here is how. Suppose $\mathbb{P}_1$ and $\mathbb{P}_2$ are two possibilities which include all the same qualitative possibilities, and yet differ with respect to the non-qualitative possibilities they include. According to Cowling (2023), we could characterise $\mathbb{P}_1$ and $\mathbb{P}_2$ by saying that they agree with respect to the distribution of qualitative properties, and disagree with respect to the distribution of non-qualitative properties. Then, if we wanted to ground further this disagreement about the distribution of non-qualitative properties, we could use haecceities, for it is easy to see that any distribution of non-qualitative properties can be defined from a distribution of qualitative and haecceitistic properties.

Notice, however, that even a nominalist about properties and relations could be a Haecceitist. For example, they could hold for independent reasons that "things could have been different non-qualitatively without being different qualitatively" (Cowling 2023) without abandoning their view that there are no such things as properties and relations. Similarly for anti-realists about haecceities. They could still claim that there are distinct maximal possibilities which agree with respect to all the qualitative possibilities they include, and yet differ with respect to the non-qualitative possibilities they

[4] Haecceities have been criticised, among others, by: Fine (1985b), Loux & Loux (1978), Menzel (2022), Moreland (2001, 2013), and Williamson (2013).

include. Most probably they would be bound to hold that these differences between possibilities are fundamental (in the sense of ungrounded, or not further explainable). However, it would be mistaken to countenance this sort of primitivism among the shortcomings of such an account — for explanations must end somewhere.

4.3 Haecceitism & PII

In what follows, I argue against the common intuition that PII-WB, a version of PII relativised to ordinary spatio-temporal entities, and Haecceitism are independent theses. In particular, I argue that if PII-WB is not necessarily true, then Haecceitism follows. This, I argue, is interesting in many respects. First, it shows that there is a connection between the identity of ordinary spatio-temporal entities and the identity of possibilities. Second, it exposes a number of authors that have denied both PII-WB and Haecceitism as holding on to an overall inconsistent position (one notable example is David Armstrong), and pressures some accounts, like Generalism, which are commonly understood to entail Anti-Haecceitism while remaining neutral on the status of PII-WB. Third, it gives the Haecceitist a very strong argument in favor of their account. There are only few authors, in fact, that still defend PII-WB as a metaphysically necessary truth.

4.3.1 Relativising PII

The Identity of Indiscernibles (PII) holds that qualitative indiscernibility is sufficient for numerical identity. We say that entities x and y are qualitatively indiscernible whenever they agree with respect to all their qualitative properties. One way to state PII is as follows:

> **Identity of Indiscernibles:** Qualitatively indiscernible entities are numerically identical.

As we saw in Section 2.2.2, by restricting the range of properties we take PII to quantify over, we obtain distinct versions of the principle. We can, for example, focus on spatio-temporal properties and understand PII as the thesis that no two entities can agree with respect to all their spatio-temporal properties. Or we can focus on intrinsic properties and interpret PII as holding that no two entities can agree with respect to all their intrinsic properties.

In line with the literature, here I take PII to be restricted to qualitative properties only, subscribing to Strawson's (1959, p. 120) motto that "[...] in the only form in which it is worth discussing, [PII holds that] it is necessarily true that there exists, for every individual, some description in purely universal, or general, terms, such that only that individual answers to that description". (As we have seen in Section 2.4, this account has been challenged in Rodriguez-Pereyra 2006.) In particular, in what follows I will focus on a version of the Identity of Indiscernibles restricted to ordinary spatio-temporal entities, which I call PII-WB:

> **PII-WB:** Qualitatively indiscernible ordinary spatio-temporal entities are numerically identical.

For a lack of better terminology, I use 'ordinary spatio-temporal entities' to indicate worldbound spatio-temporal beings (like tables and chairs) which are neither worlds nor possibilities.[5]

4.3.2 The Common Intuition

Now, it is commonly held that PII-WB is independent from Haecceitism, in the sense that both PII-WB and its negation are compatible with both Haecceitism and Anti-Haecceitism. This is because, arguably, facts about the identity of ordinary spatio-temporal entities do not have direct bearing on facts about the identity of possibilities. To explain. Suppose PII-WB is necessarily true. It follows that in the actual world there are no two ordinary spatio-temporal entities which share all their qualitative properties. In particular, you and I differ with respect to at least some qualitative property. This alone doesn't seem to commit you to the fact that I could have had all the qualitative properties you actually have and you could have had all the qualitative properties I actually have (Haecceitism), nor to the fact that I could have never had all the qualitative properties you actually have and you could have never had all the qualitative properties I actually have (Anti-Haecceitism). Similarly, suppose PII-WB is not necessarily true. Then there is a world where there are two indiscernible but distinct entities. For vividness, take Lewis's (1986, pp. 230–231) world with two qualitatively indiscernible twins. Are you bound, by the existence of this world alone, to hold that the indiscernible twins could have swapped their qualitative role? It doesn't seem so. Lewis famously claims it is plausible that his twins could have swapped their role: however, he gives no argument for why this is the

[5] The qualification 'ordinary' is meant to exclude black holes, quarks, etc. as well as alien *possibilia*.

case. He relies on his intuitions, which happen to be in line with Haecceitism. However, there is nothing, on the face of it, that prevents you from holding the contrary intuition: that it would have been impossible for the twins to swap their qualitative roles.

4.3.3 Lewis on Haecceitism

It is interesting to note that although in this passage Lewis (1986, pp. 230–231) seems to endorse Possibilities Haecceitism, he is definitely not a friend of Possible Worlds Hecceitism. (See Section 4.2.1 and Section 4.2.2.) This is all the more interesting, given that according to Lewis's Modal Realism, for any way the actual world could be/could have been, there is a possible world which represents such possibility.

It is through a clever use of his famous as well as controversial counterpart theory, that Lewis can maintain that Possible Worlds Haecceitism is false whilst endorsing Possibilities Haecceitism. Here is how. Recall that, according to Lewis's Modal Realism, *de re* modal statements such as 'Napoleon could have won the battle of Waterloo' must be analysed in terms of counterparts. That is: the sentence 'Napoleon could have won the battle of Waterloo' is true if and only if there is a possible world w inhabited by a counterpart of Napoleon who happened to win the battle of Waterloo in w.

Recall further that some entity x in a world w is a counterpart of Napoleon if x is *similar* to Napoleon in some relevant respects, where 'similar', in Lewis's account, means 'qualitatively similar'. This, Cowling (2023) remarks, entails that qualitative indiscernible worlds, in Lewis's Modal Realism, always represent the same *de re* possibilities. *A fortiori*, this entails that Possible Worlds Haecceitism, as defined in Section 4.2.2, is false.

What makes it possible for Lewis to endorse Possibilities Haecceitism whilst rejecting Possible Worlds Haecceitism is that, in Lewis's account, individuals can have counterparts which inhabit their same possible world. This is how Lewis manages to represent haecceitistic differences between possibilities. One and the same world, for Lewis, can represent distinct possibilities whenever it contains one or more individuals standing in the counterpart relation to each other.

Consider again the world, mentioned in Section 4.3.2, with two qualitatively indiscernible twins (Lewis 1986, pp. 230–231). Since the twins are each other's counterparts, this world represents two possibilities at once. And since they are represented by one and the same world, these possibilities must include all the same qualitative possibilities. But then they differ only with respect to the non-qualitative possibilities they include, thus entailing Possibilities Haecceitism.

4.4 An Argument from Philosophy of Physics

In this Section, I quickly discuss a famous argument in Philosophy of Physics which aims to establish that indistinguishable particles are non-individual entities, on pain of contradiction. The ideas behind this argument will play a major role in the argument I will present in Section 4.5.

4.4.1 Preliminaries: Individuals vs. Non-individuals

Non-individuals are entities to which identity does not apply. If x is a non-individual, then sentences like 'x is self-identical' and 'x is not identical to itself' are meaningless. Similarly, if x and y are non-individuals, then sentences like 'x is identical to y' and 'x is distinct from y' are meaningless too. More formally, we say that if 'x' refers to a non-individual, then 'x' cannot meaningfully flank any identity symbol.[6]

The following argument shows how we can arrive at the conclusion that some entities are non-individuals by considering the relations between two scenarios which differ solely with respect to *which* entity is *which*. The argument goes as follows.

4.4.2 The Argument

Let $\mathbb{C}1$ be a configuration with only two indistinguishable elementary particles x and y in different energy states: $\mathbb{E}1$ and $\mathbb{E}2$ respectively.[7] And let $\mathbb{C}2$ be a configuration disagreeing with $\mathbb{C}1$ only with respect to which particle is in $\mathbb{E}1$ and which particle is in $\mathbb{E}2$. We say that $\mathbb{C}2$ is a permutation of $\mathbb{C}1$. Now, either (1) $\mathbb{C}2$ is the same as $\mathbb{C}1$, or (2) $\mathbb{C}2$ is distinct from $\mathbb{C}1$.

Since by assumption x and y are indistinguishable, they are subject to the so-called Indistinguishability Postulate, according to which there is no way, even in principle, to distinguish states that differ only by a permutation

[6]For more about non-individuality in Quantum Mechanics, see Landau & Lifschitz (1959), and Post (1963). I will discuss in more details whether a term can singularly refer to a non-individual entity in Chapter 8.

[7]Examples like this are common in Quantum Mechanics. See, among others, French (1989), French & Krause (2006), Saunders (2003), and Berto (2017).

of their (relevant) particles.

According to the Received View of Quantum Mechanics, championed among others by French (1989, p. 154), the Indistinguishability Postulate entails that $\mathbb{C}1$ and $\mathbb{C}2$ are the same configuration, and so it rules out (2). Suppose then that x and y are individuals, that is: entities to which identity applies. It follows that $\mathbb{C}1$ satisfies the following sentence:

"The particle in state $\mathbb{E}1$ is distinct from the particle in state $\mathbb{E}2$",

where the expression "the particle in state $\mathbb{E}1$" is understood *de re* (or as a referential description).[8] However, given that $\mathbb{C}1 = \mathbb{C}2$, $\mathbb{C}1$ also satisfies the sentence:

"The particle in state $\mathbb{E}1$ is not distinct from the particle in state $\mathbb{E}2$".

This is because, understood *de re*, the expression "The particle in state $\mathbb{E}1$" refers to the same individual in both sentences. But by assumption $\mathbb{C}2$ is a permutation of $\mathbb{C}1$, therefore the particle that is in $\mathbb{E}1$ according to $\mathbb{C}1$ is in $\mathbb{E}2$ according to $\mathbb{C}2$. It follows that $\mathbb{C}1$ satisfies an inconsistent set of sentences, which is a contradiction. It follows that identity must not apply to x and y, namely: x and y are non-individuals.

4.4.3 The Received View of Quantum Mechanics

The above argument concludes that x and y are non-individuals by contradiction from the assumption that $\mathbb{C}1$ and $\mathbb{C}2$ are the same configuration. In

[8] See Donnellan (1966).

the next Section, I will use a similar reasoning to show that if PII-WB is not necessarily true, then Haecceitism is true.

However, it is important to note that the argument just rehearsed heavily relies on a peculiar interpretation of the Indistinguishability Postulate, according to which it entails that $\mathbb{C}1$ and $\mathbb{C}2$ are one and the same configuration. And witness of the fact that this interpretation can be challenged is the fact that although widespread, the Received View of Quantum Mechanics is not unanimously accepted. Many authors remain who believe quantum particles are individuals, in the sense of having determinate identity conditions. (See, among others, Bueno 2014, and Jantzen 2011.)

This is the reason why I will not make use of the Indistinguishability Postulate. I aim, in fact, at a conclusion that is general enough to apply not only to elementary particles, and which is independent from the consistency and tenability of the Received View of Quantum Mechanics.

4.5 From PII to Haecceitism

Recall PII-WB is the thesis that no two distinct ordinary entities can agree with respect to all their qualitative properties. PII-WB can be regimented in a second-order language as:

$$\forall x \forall y (\forall P (Px \leftrightarrow Py) \rightarrow x = y)$$

with 'x' and 'y' individual variables ranging over ordinary spatio-temporal entities only, and 'P' a second-order predicate variable ranging over qualitative properties. We then say that PII-WB is false if and only if the sentence:

$$\exists x \exists y (\forall P(Px \leftrightarrow Py) \land x \neq y)$$

is true.[9] We read this sentence as: "There are at least two ordinary spatio-temporal entities which are qualitatively indiscernible and yet distinct". Therefore, we say that PII-WB is not a necessary truth if and only if there is a possibility where this last formula is true. It is worth stressing this last point. It tells us that in order for a possibility $\mathbb{P}$ to violate PII-WB, the indiscernible ordinary entities x and y in $\mathbb{P}$ must satisfy the formula '$x \neq y$', which means that x and y cannot be non-individuals. If they were, in fact, 'x' and 'y' wouldn't be able to flank the identity symbol in '$x \neq y$', and this would make the relevant formula unsatisfiable.

4.5.1 The Argument

I want now to show that if PII-WB is not necessarily true then Haecceitism is true. So let us start by assuming that PII-WB is not necessarily true. It follows that there is a non-empty class of possibilities according to which there are at least two ordinary spatio-temporal entities which are qualitatively indiscernible and yet distinct. Let $\mathbb{P}1$ be one such possibility, and call two of the indiscernible entities in $\mathbb{P}1$ Adam and Beth.

I assume size is a contingent property of Adam and Beth, and I hold this is an innocent assumption. All the ordinary spatio-temporal entities I can think of have their size contingently: I could have been taller, you could have been shorter — and this seems to hold for ordinary spatio-temporal entities generally. Then, there is a possibility $\mathbb{P}2$ according to which Adam and Beth

[9] Where again, 'x' and 'y' range over ordinary spatio-temporal entities and 'P' ranges over qualitative properties only.

are qualitatively indiscernible and yet distinct, and their size is such that they do not occupy all the space available.[10]

Since Adam and Beth are spatio-temporal entities, then they must be at some distance from each other. A (partial) representation of $\mathbb{P}2$ is the following:

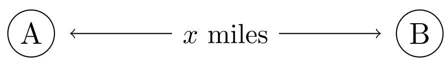


Figure 4.1: A representation of $\mathbb{P}2$.

According to $\mathbb{P}2$, Adam and Beth occupy only a small portion of the space available. Therefore, I hold that if $\mathbb{P}2$ is a genuine possibility, then Adam and Beth could have existed alongside another couple of indiscernible spatio-temporal entities, call them Charlie and Dave, and that (1) the distance between Adam and Charlie was the same as the distance between Beth and Dave, (2) the distance between Adam and Beth was different from the distance between Adam and Charlie, (3) the distance between Charlie and Beth was different from the distance between Beth and Adam, and (4) the distance

[10]One might object that Adam and Beth having their size contingently is not enough to conclude that they could have existed and occupied only a small portion of the space available, for it might be the case that it is essential to both Adam and Beth that they can coexist with another qualitatively indiscernible entity only if together with it they occupy the entirety of the available space. This is an interesting objection, but one I find very implausible. First, remember that Adam and Beth are by definition ordinary spatio-temporal entities, and ordinary spatio-temporal entities usually don't have such relational essential spatial properties. So why Adam and Beth should have them? This asymmetry should be explained, and I cannot think of any plausible and non *ad hoc* explanation one could give. Second, remember that $\mathbb{P}1$ is an arbitrary counterexample to PII-WB. Therefore, holding that Adam and Beth have such essential spatial properties entails that there cannot be any counterexamples to PII-WB where the entire space is not fully occupied. This, however, flies in the face of the virtually unanimous consensus that alleged counterexamples to PII-WB include Black's (1952) indiscernible spheres. For these reasons, I hold this objection doesn't go through.

between Adam and Charlie was different from the distance between Charlie and Beth. We can represent this possibility, which we call $\mathbb{P}3$, as follows:

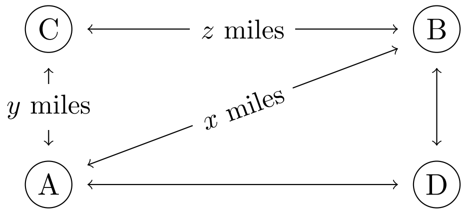


Figure 4.2: A representation of $\mathbb{P}3$.

One way $\mathbb{P}3$ is able to discern between Adam and Beth is by means of the non-qualitative property 'being y miles away from Charlie', which only Adam instantiates. Therefore, according to $\mathbb{P}3$ there are four qualitatively indiscernible objects and a non-qualitative property which doesn't depend either on the identity of Adam or on the identity of Beth and discerns Adam from Beth insofar as only one of them instantiates it.

From this we conclude that, if PII-WB is not necessarily true, then there is a possibility according to which there are at least two indiscernible objects and there is a non-qualitative property which discerns them and doesn't depend on neither of their identities.

Now we can ask whether it could have been the case that it was Dave, and not Charlie, that was y miles away from Adam. (Remember, we are still working under the assumption that PII-WB is not necessarily true.) Suppose so. Then we have $\mathbb{P}4$, according to which there are four indiscernible spatio-temporal entities: Adam, Beth, Charlie, and Dave, and the distance between Adam and Dave is of y miles, while the distance between Adam and Charlie is of z miles. (The distance between Adam and Beth is still of x miles.)

A (partial) representation of $\mathbb{P}4$ is the following:

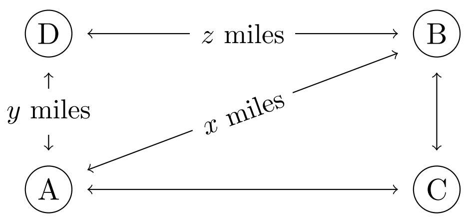


Figure 4.3: A representation of $\mathbb{P}4$.

Is $\mathbb{P}4$ possible? I think it is, for two reasons. First, I hold the differences between $\mathbb{P}3$ and $\mathbb{P}4$ are not enough to make one of them an impossibility. After all, $\mathbb{P}3$ and $\mathbb{P}4$ have the same number of entities, with the same intrinsic properties, in exactly the same spatial configuration. They only differ with respect to *which* entity is y miles away from Adam and *which* entity is y miles away from Beth — and I hold this cannot be enough reason to say that $\mathbb{P}4$ is impossible. For then, what would make $\mathbb{P}3$ possible? If you think that $\mathbb{P}3$ is possible and $\mathbb{P}4$ is not you are committing yourself to the thesis that some spatial properties are essential to some (and only some) of our indiscernible entities. After all, you are saying that other things being equal, Dave could not have been in another part of the universe, not already occupied by any other entity — and this seems wrong to me. Note further that to say that $\mathbb{P}4$ is possible is not equivalent to a commitment to Haecceitism: for there is no reason, yet, why we should think that $\mathbb{P}3$ and $\mathbb{P}4$ are distinct possibilities.

Second, consider again $\mathbb{P}3$. I hold that it is not *necessary* for the entities in $\mathbb{P}3$ that they are at rest. That is: we can safely assume that, being ordinary spatio-temporal entities, Adam, Beth, Charlie and Dave have their position in space only contingently.[11] Then, it is possible that Charlie and Dave could have been moving around Adam and Beth, at the same speed, along the same orbit. Call this possibility ℙ3.5:

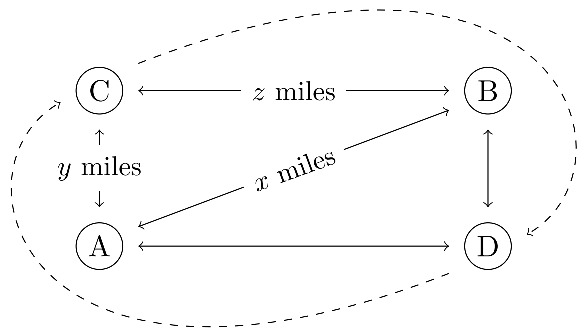


Figure 4.4: A representation of ℙ3.5.

If ℙ3.5 is possible then for any spatial configuration our spatio-temporal entities are in at some moment of their revolution, there is a possibility such that (1) they are in exactly the same configuration, and (2) they will never be and have never been in any other configurations. (Notice that ℙ3.5 doesn't need to be *nomologically* possible, only *metaphysically* possible.) In particular: at any moment of Charlie's revolution it is possible that Charlie could have been at that same distance with respect to all the other entities in the world, without having never moved. (The same holds for Dave.) If this is the case, then we find ℙ4 among the possibilities generated from ℙ3.5. And since I see no way to deny that Charlie and Dave could have been unmoving

[11] Like the assumption on the contingency of the size of ordinary spatio-temporal entities, I hold that the supposition that ordinary spatio-temporal entities have their spatial position contingently is philosophically innocuous. After all, for example, the Earth could have been closer to the Sun than it actually is.

(notice that to deny this would be to deny that $\mathbb{P}3.5$ is possible), then it seems that indeed if $\mathbb{P}3$ is possible, then so must be $\mathbb{P}4$.

We conclude that $\mathbb{P}4$ is possible if $\mathbb{P}3$ is, and that $\mathbb{P}3$ is possible if PII-WB is not necessarily true (which we have assumed at the start). Now we have two cases: either (1) $\mathbb{P}3$ is the same as $\mathbb{P}4$, or (2) $\mathbb{P}3$ is distinct from $\mathbb{P}4$.

A quick moment's thought at $\mathbb{P}3$ and $\mathbb{P}4$ reveals that they are both *maximal* and *include the same qualitative possibilities*. They in fact represent total ways the world could have been, and agree with respect to all the qualitative possibilities they include: the possibility *that some spatio-temporal entity is x miles away from some other spatio-temporal entity*, the possibility *that some spatio-temporal entity is y miles away from some other spatio-temporal entity*, the possibility *that all entities are extended*, etc. This means that if Anti-Haecceitism is true, then (1) is true.

So assume for *contradiction* that Anti-Haecceitism is true, that is: that (1) holds and $\mathbb{P}3$ and $\mathbb{P}4$ are one and the same possibility.[12] Suppose further that Adam, Beth, Charlie and Dave are individuals: that is, entities to which identity applies. Then, $\mathbb{P}3$ satisfies the sentence:

"The spatio-temporal entity which is y miles away from Charlie is distinct from the one which is y miles away from Dave",

where the expression "the spatio-temporal entity which is y miles away from Charlie" is understood *de re* (or as a referential description). However, since $\mathbb{P}3 = \mathbb{P}4$, then $\mathbb{P}3$ also satisfies the sentence:

[12]The assumption that Anti-Haecceitism is true is doing the same work for us that the Indistinguishability Postulate does for the Received View of Quantum Mechanics in the argument discussed in Section 4.4.2.

"The spatio-temporal entity which is y miles away from Charlie is not distinct from the one which is y miles away from Dave".

This is because, understood *de re*, the expression "the spatio-temporal entity which is y miles away from Charlie" refers to the same individual in both sentences — and since $\mathbb{P}4$ is nothing more than a permutation of $\mathbb{P}3$, the entity which is y miles away from Charlie according to $\mathbb{P}3$ is y miles away from Dave according to $\mathbb{P}4$.

Since $\mathbb{P}3$ satisfies an inconsistent set of sentences it cannot be a possibility, and we have a contradiction. To avoid this, we must conclude that one of our assumptions is false — and since we are reasoning under three assumptions, we only have three ways out of inconsistency. The first is to reject Anti-Haecceitism, which leads us to the conclusion that if PII-WB is not necessarily true, then Haecceitism follows.[13] The second is to reject the assumption that PII-WB is not necessarily true. This allows us to conclude that if Anti-Haecceitism holds, then PII-WB is necessarily true — which is equivalent, by contraposition, to the claim that if PII-WB is not necessarily true, then Haecceitism follows.

The last possibility is to reject the assumption that Adam, Beth, Charlie and Dave are indeed individuals. This amounts to say that at least one of our entities is a non-individual, in the sense specified in Section 4.4.1. However, since the property 'being a non-individual' doesn't depend on the identity of any specific individual, then if one of our entities has it, all of them must have it — for they are, after all, qualitatively indiscernible. But if this is the

[13]Note that this move is equivalent to rejecting case (1) and accepting (2), according to which $\mathbb{P}3$ and $\mathbb{P}4$ are distinct possibilities. But then, since $\mathbb{P}3$ and $\mathbb{P}4$ are maximal and include the same qualitative possibilities, we have Haecceitism.

case, then Adam and Beth are non-individuals, and $\mathbb{P}1$ is not a possibility in which we have two *distinct*, in the sense of non-identical, spatio-temporal entities. Therefore, $\mathbb{P}1$ is not a counterexample to PII-WB anymore, for it doesn't satisfy the conjunction:

$$\exists x \exists y (\forall P(Px \leftrightarrow Py) \wedge x \neq y).$$

This is because, since identity does not apply to Adam and Beth, we cannot obtain the second conjunct: '$x \neq y$'. Therefore, since $\mathbb{P}1$ was chosen arbitrarily, then PII-WB is necessarily true, which contradicts our first assumption. By contradiction, this route leads us once again to the conclusion that if PII-WB is not necessarily true, then Haecceitism is true.

4.6 Philosophical Discussion

The argument I presented (henceforth I call it NPH for short) shows that Haecceitism follows from the negation of the necessary truth of PII-WB. This is interesting for a number of reasons.

4.6.1 PII & Transworld Identity

First, NPH makes clear that there is a direct connection between how we consider the identity of worldbound individuals and the number of genuine possibilities. To explain. In the literature on Haecceitism, different versions of the Identity of Indiscernibles are distinguished. One option, which we can call PII-T (for PII-Transworld), holds that no two individuals in the entire logical space can agree with respect to all their qualitative properties. PII-T either rules out the possibility of there being distinct worlds which contain indiscernible individuals, or bounds us to some account of transworld identity. This is because, if PII-T is true and we hold that distinct worlds w and v could contain indiscernible individuals x and y, then PII-T entails that w and v overlap, for it entails that x is the same as y. (And since w and v overlap, we have transworld identity.) It is therefore not surprising that PII-T is not independent from Haecceitism. In the end, Haecceitism is a thesis about transworld identity.[14]

PII-WB, on the other hand, only claims that no two individuals *in a given world* can be qualitatively indiscernible. In this sense, it is a weaker thesis than PII-T. Furthermore, on the face of it, PII-WB doesn't seem to entail any fact about transworld identity. This is why it is so surprising that this version of PII, or better, its negation, entails Haecceitism, which is, once again, a thesis about transworld identity. This tells us that the number of indiscernible entities we take to exist at a given world determines a lower bound for the number of possibilities we have to include in our metaphysics.

4.6.2 Pressuring Extant Philosophical Theories

Second, NPH puts pressure on some authors and extant philosophical theories. The literature on Haecceitism and the Identity of Indscernibles reveals a difference in number between the authors who believe that PII-WB is not necessarily true and the number of authors who endorse Haecceitism.[15] Al-

[14]For more on this, see Mackie (2006).

[15]On the Identity of Indiscernibles, see: Adams (1979), French & Redhead (1988), and Hawley (2009). On Haecceitism, see: Cowling (2012), Kment (2012), Plantinga (1974), and Skow (2008).

though the majority of authors, in fact, seem to find counterexamples to PII-WB strikingly convincing, Haecceitism seems still to be a position that few are willing to explicitly endorse. If NPH is sound, however, then all authors who believe PII-WB to be possibly false are indeed committed to Haecceitism. The fact that at the moment this proportion is not met can clearly be explained by noticing that many authors have explicitly worked on only one of these theses and, by not recognising the intimate connection between the two, didn't explicitly endorse a position on the latter based on their position on the former.

However, one can also find authors that have explicitly denied both PII-WB and Haecceitism, therefore holding on to an overall inconsistent position. One such author is David Armstrong. In his *Universals and Scientific Realism, Vol.1*, Armstrong argues against the Bundle Theory of Substance on the basis of its commitment to PII-WB (Armstrong 1978, ch. 9). He suggests that since PII-WB is false, and since the Bundle Theory entails PII-WB, then one must reject the Bundle Theory as being false too. His denial of PII-WB should make Armstrong a Haecceitist. However, Armstrong himself (1989, pp. 57–61) specifically endorses a version of Anti-Haecceitism. Another such author is Thomas Hofweber, who flirts with Anti-Haecceitism despite remaining officially neutral about Haecceitism (Hofweber 2005, p. 27), and yet accepts that PII-WB is not necessarily true (Hofweber 2015, p. 476).

Furthermore, NPH puts pressure on some current philosophical theories, among which we find Generalism, as defended in Dasgupta (2009) and Turner (2016). According to Generalism, there are no primitive individuals. The structure of reality is instead taken to be fundamentally general. This is best explained in terms of qualitative properties. Reality, the Generalist argues, is exhausted by facts about the distribution of qualitative properties; facts about individuals are not required. Cowling (2023) argues that insofar as Haecceitism presupposes distinct maximal possibilities which only differ with respect to the identity of individuals (and those non-qualitative properties that this difference in identity entails), Generalism rules out Haecceitism. However, Dasgupta (2009, p. 49) argues that Generalism is compatible with PII-WB being possibly false. But if so, Generalism turns out to be an inconsistent view.

Another such theory is Necessitarianism, according to which the only maximal possibility out there is the actual one. Necessitarianism contradicts Haecceitism, for if there is only one maximal possibility, then no two distinct possibilities can disagree with respect to all the non-qualitative possibilities they include. However, Necessitarianism entails that all truth are necessary, and therefore, to avoid inconsistency, Necessitarianists must hold, *contra* French (1989), that PII-WB is necessarily true.

4.6.3 An Argument for Haecceitism

Finally, NPH is an argument in favor of Haecceitism, on the assumption that PII-WB is not a necessary truth. At present, the main arguments for Haecceitism are all arguments from conceivability. Cowling (2023) suggests these arguments need two steps: a *conceivability step*, where it is argued that some scenario S is conceivable, and a *possibility step*, where it is argued that since S is conceivable, then S is possible. Therefore, Cowling (2023) remarks, there are multiple ways for the Anti-Haecceitist to challenge any of those arguments. I think NPH represents a novel strategy for the Haecceitist, and suggest that despite it still involves some form of conceivability (in the sense that almost all arguments against PII-WB require some conceivability step), it is stronger than all other arguments at the Haecceitist's disposal. To see why this is the case, take one of the most influential argument for Haecceitism from the possibility of Max Black's (1952) scenario against PII-WB. The argument, as reconstructed by Cowling (2023), is the following:

P1: Black's (1952) universe, containing nothing more than two indiscernible spheres, is a genuine possibility.

P2: If Black's universe is a genuine possibility, then we can conceive of a world containing only two indiscernible spheres.

P3: If we can conceive of a world containing only two indiscernible spheres, then we can conceive of distinct worlds that differ only insofar as these spheres swap their qualitative role.

P4: If we can conceive of distinct worlds that differ only insofar as the relevant spheres swap their qualitative role, then these distinct world are possible.

C: Therefore, Haecceitism follows.

We can see that there is plenty of premises the Anti-Haecceitist can challenge. In particular, **P4** is very weak, for many would argue that conceivability is indeed no good guide to possibility.[16] NPH, on the other hand, only needs **P1** to conclude that Haecceitism is true.

Clearly, the Anti-Haecceitist might yet be arguing against **P1**. However, notice that this strategy is going to affect NPH as well as the conceivability argument just laid down. Also, to argue against **P1** in all the versions this argument could come is tantamount to argue that PII-WB is indeed necessarily true.

Furthermore, by establishing that if PII-WB is not necessarily true then Haecceitism follows, NPH also establishes that Anti-Haecceitism entails that PII-WB is a necessary truth. And this, I suggest, doesn't look good for the Anti-Haecceitist. As I mentioned before, few authors remain who endorse PII-WB as a necessary truth, and if Anti-Haecceitism entails PII-WB, than the same argument which Armstrong and others have used against the Bundle Theory of Substance can be used *mutatis mutandis* against Anti-Haecceitism.

4.6.4 Weak Haecceitism

I conclude by noticing that although the denial of PII-WB commits one to Haecceitism, it doesn't bound them to accepting *all* haecceitistically distinct possibilities. That is: one can deny PII-WB and therefore endorse Haecceitism without being committed to possibilities like the one I described in the Section 4.2, where you and I swapped our qualitative role. In fact, holding that PII-WB is not necessarily true entails a very weak version of Haecceitism, that is:

[16] See Yablo (1993).

Weak Haecceitism: There are maximal possibilities which include all the same qualitative possibilities, and yet differ with respect to the non-qualitative possibilities they include, these last possibilities concerning only qualitatively indiscernible individuals.

And since denying PII-WB entails Weak Haecceitism, it entails Haecceitism *a fortiori*. However, it doesn't commit us to any possibilities in which two qualitatively discernible individuals, like you and me, swap their qualitative role.

4.7 Conclusion

In this Chapter I argued that if a popular version of the Identity of Indiscernibles relativised to ordinary spatio-temporal entities (PII-WB) is not necessarily true, then Haecceitism follows — where Haecceitism is the view that there are maximal possibilities which include all the same qualitative possibilities, and yet differ with respect to the non-qualitative possibilities they include.

I argued that this is a strong result in favor of Haecceitism, for the majority of authors today find counterexamples to the Identity of Indiscernibles extremely convincing, and many philosophical positions have been attacked and continue being attacked on the basis of their commitment to the Identity of Indiscernibles. Also, I put some pressure on some authors and current philosophical accounts, by showing that they are holding on to inconsistent claims. Finally, since I take my argument to be independent from any particular account of possible worlds/possibilia, I take it to show that all these accounts must share a connection between the number of ordinary spatio-temporal entities within given possibilities and the number of overall possibilities.

Chapter 5

On Qualitative Properties

As we have seen in the preceding Chapters, qualitative properties play an important role in debates about PII and indiscernibility. Introduced almost ninety years ago by Carnap (1937, p. 45), the distinction between qualitative and non-qualitative properties is of considerable significance in philosophy, being essential in the analysis of laws and explanation, Haecceitism, and Physicalism (just to name a few).

Notwithstanding, the jury is still out on how to distinguish between qualitative and non-qualitative properties across the board. Virtually everyone agrees that properties like 'being red' are qualitative, and properties like 'being identical to Donald Trump' are not. Outside a handful of properties, however, there is still much disagreement. Is the property 'being an actual donkey' qualitative? And what about the property 'being even'? As we will see below, intuitions about these and similar properties are still a matter of contention, with some authors placing the property 'being an actual donkey', for example, among the qualitative properties (see, among others, Rosenkrantz 1979), while others arguing that this and similar properties should be understood as non-qualitative (see, among others, Cowling 2015). And to further complicate the matter, to date there is no agreed upon definition of qualitative properties on the market. As we will see in Section 5.2, in fact, all the attempts advanced so far at making the qualitative/non-qualitative distinction theoretically precise have been found wanting, in one way or another.

In previous Chapters I have characterised qualitative properties rather informally, as those properties which do not depend on the identity of any specific individuals. Here I aim to set up a new working account of qualitative and non-qualitative properties. I start by discussing the role that qualitative properties plays in the definition of many philosophical notions (Section 5.1), and by reviewing some of the most famous attempts at defining qualitative properties (Section 5.2). I then present my new account of qualitative properties (Section 5.3), and develop a formal framework inspired by such an account (Section 5.4). Finally, I the framework against different kinds of properties (Sections 5.5 and 5.6). I argue that my account is able to provide new valuable intuitions when it comes to properties whose status is still a matter of debate, while at the same time it aligns with the main intuitions with respect to those properties about which virtually everyone agrees.

5.1 Qualitative Properties in Philosophy

Many authors agree that it is of paramount importance for contemporary philosophy to find a way to make the distinction between qualitative and non-qualitative properties theoretically precise. (See, among others, Cowling 2015, p. 278–281 and Hoffmann-Kolss 2019, p. 996.)

This is because qualitative and non-qualitative properties play an essential role in many philosophical debates, spanning from Haecceitism, to supervenience, to accounts of laws and explanations. The idea is then that a precise understanding of these debates requires a precise understanding of what makes some properties qualitative and others not.

For example. Some metaphysical accounts of laws and explanations dictate that only qualitative properties can feature in law-like statements. (See, among others, Hempel & Oppenheim 1948.) Similarly, whether determinism is incompatible with alternative non-qualitative possibilities is an issue that can only be settled once a sufficiently good understanding of what it means for a property to be qualitative is reached. (See, among others, Brighouse 1997, and Melia 1999.)

Akin is the situation with theses like Haecceitism, which we discussed at length in Chapter 4. As the reader will remember, Haecceitism is the thesis that there are distinct possibilities which only differ in so far as they represent different non-qualitative facts. What Haecceitism amounts to, and the repercussion it has for the Philosophy of Modality if true, are all functions of what we mean when we say that a property is qualitative.

The distinction between qualitative and non-qualitative properties is also important for defining intrinsic properties. Among others, in fact, Langton and Lewis (1998) believe that all intrinsic properties are qualitative. And, Marshall (2018) notes, the notion of intrinsic properties is itself at the heart of many philosophical theses. It is widely employed, for instance, in distinguishing 'real' changes from mere 'Cambridge' changes (as per Geach 2000), and is one of the main pieces of artillery in Lewis's (1986; 1988) argument against three-dimensionalism. Intrinsic properties are also employed in many definitions of the notions of supervenience, for example in Jackson (1998), and Kim (1982). As Hoffmann-Kolss (2019, p. 996) nicely explains, the thought behind restricting definitions of supervenience to qualitative properties is that there are some non-qualitative properties that cannot be instantiated by more than one individual. (Think, for example, at the property 'being identical to Napoleon'.) Many definitions of supervenience have the same form: given a type X and a type Y of properties, we say that properties of type X supervene on properties of type Y if and only if any two individuals which are indiscernible relative to properties of type Y are also indiscernible relative to properties of type X. Notice however that if non-qualitative properties which cannot be instantiated by multiple distinct individuals are included in type Y, then the any definition of this kind would be rendered trivial, for there would be no two individuals which are indiscernible relative to properties of type Y. (See also Bennett 2004, and Horgan 1982.)

Qualitative properties heavily feature also in debates about Physicalism, where it is often argued that any distribution of properties supervenes on the distribution of some physical properties, which are in turn thought to be inherently qualitative. (See, for example, Chalmers 1996, and Daly & Liggins 2010.)

Finally, Cowling (2015) suggests that qualitative properties also play a substantive role in the Philosophy of Language. One notable example is the case of descriptivism, which holds that proper names are semantically equivalent to definite descriptions. Descriptivism has faced numerous challenges, among which we find Putnam's (1975) Twin Earth, which shows, according to Cowling, that proper names can be used to refer to only one of many qualitatively indiscernible entities. Cowling (2015, p. 279) remarks that Putnam's thought experiment is a serious threat to descriptivist positions only insofar as descriptivists limit themselves to definite descriptions containing only predicates referring to qualitative properties.

5.2 Defining Non-Qualitative Properties

Many authors have attempted to define the notions of qualitative and non-qualitative properties. Some of them have have attempted to ground the distinction between qualitative and non-qualitative properties in more fundamental notions, like the notion of 'individual', while others have understood the distinction as not amenable of further reduction, usually taking haecceities as primitive and paradigmatic cases of non-qualitative properties, and building the distinction up from there.

Non-qualitative properties are commonly characterized as those properties which *depend*, in some way or other, on specific individuals. Different accounts of qualitative properties will then usually correspond to different notions of *dependence* and different *relata* of the dependence relation.

5.2.1 Adams's Definition

As we have already mentioned in Section 2.2.3, one of the most interesting definition of qualitative properties comes from Adams (1979). Here, the author defines qualitative properties on the basis of how they can be expressed. In particular, Adams (1979, p. 7) writes:

> [...] a property is purely qualitative [...] if and only if it could be expressed, in a language sufficiently rich, without the aid of such referential devices as proper names, proper adjectives and verbs (such as 'Leibnizian' and 'pegasizes'), indexical expressions, and referential uses of definite descriptions.[1]

It is important to notice that Adams is not defining qualitative properties on the basis of the syntactic content of the predicates we use (or might want to use) to regiment them. Indeed, he is defining qualitative properties on the basis of the syntactic content of all the predicates that *could possibly be used* to regiment them. In other words: there is no specific language Adams has in mind in which to evaluate the expressions referring to qualitative or non-qualitative properties. Instead, he is quantifying over the set of all possible (sufficiently rich) languages.

Although at first sight it may look like a strength, its incredible degree of generality turns out to be this definition's greatest weakness. It paves the way, in fact, for a knock-down argument proposed by Cowling (2015), which

[1] Just after this passage, Adams (1979, p. 7–8) suggests another definition of qualitative properties, according to which a property P is qualitative if and only if P is either a 'basic suchness', or P is constructed out of basic suchnesses by means of logical operations. For Adams, a property P is a basic suchness if and only if: (1) P is not an haecceity and it is not equivalent to any haecceity, (2) P is not a property of being related in some way or another to some individual, and (3) "[P] is not a property of being identical with or related in one way or another to an extensionally defined set that has an individual among its members, or among its members' members, or among its members' members' members, etc." Although Adams (1979, p. 7) remarks that this second definition of qualitative properties is more illuminating than his first one, the literature almost uniquely focused on Adams's first definition.

rests on two pillars: a charge of incompleteness and a charge of confusion. According to Cowling (2015, p. 287), Adams's definition is inherently incomplete, and necessarily so. How could we, in fact, even begin to enumerate all possible linguistic types which might be associated, *in all possible languages*, to non-qualitative properties?

Furthermore, Cowling (2015) remarks, this definition confuses the order of explanation. If indeed (at least some) qualitative properties are not dependent on the mind, then it should be the fact that certain properties are qualitative which explains why we use certain expressions instead of others, and not the other way around. This is because language is mind-dependent, and therefore the overall definition rests on an implausible assumption about the concordance between the reality we inhabit and the language we use to describe it.

Cowling's argument generalises widely, for it applies to all accounts of qualitative properties which ground the distinction between qualitative and non-qualitative properties in facts about linguistic expressions and linguistic types.

5.2.2 Rosenkrantz's Definition

Another attempt at defining qualitative properties comes from Rosenkrantz. According to Rosenkrantz (1979, p. 518), non-qualitative properties are those properties which have individuals as concrete constituents.

Rosenkrantz's definition of a concrete constituent is the following. Given a property P and an individual x, we say that x is a concrete constituent of P if and only if: (1) x is a contingently existing and concrete individual, (2) P is possibly instantiated, and (3) P is necessarily such that, if P is instantiated, then x exists at some time. (Notice the *de re* modalities in the last two conditions.)

For example, think about the property 'standing next to Joe Biden'. Joe Biden is clearly a contingently existing and concrete individual. Furthermore, the property 'standing next to Joe Biden' is possibly instantiated, which means that there is a possible world where this property is instantiated. We can easily imagine that, if there is no one now in the actual world next to Joe Biden, someone could still be. This means that there is at least one possible world where there is someone which is standing next to Joe Biden, and hence our property is possibly instantiated. Finally, if the property 'standing next to Joe Biden' is instantiated, then there is some time in which Joe Biden exists. (For if he didn't exist, how could someone be standing next to him?) It would be impossible, in fact, for the property 'standing next to Joe Biden' to be instantiated were Joe Biden never to have existed. Therefore, we can say with Rosenkrantz that the property 'standing next to Joe Biden' has Joe Biden as concrete constituent. It follows that, according to Rosenkrantz's definition, 'standing next to Joe Biden' is non-qualitative.

Rosenkrantz's definition correctly characterises properties like 'being red' and 'being (identical to) Donald Trump'. However, according to Hoffmann-Kolss (2019, p. 1000–1001), it mischaracterises properties like 'being Donald Trump's counterpart', and 'hallucinating the Eiffel Tower': although these two properties are in fact intuitively non-qualitative, "[their] instantiation by some individual at some possible world w is compatible both with the existence and the non-existence of [Donald Trump/the Eiffel Tower] at w" (Hoffmann-Kolss 2019, p. 1001).

Rosenkrantz's definition is also challenged by Cowling (2015, p. 288) on the basis of its inability to correctly characterise negative haecceities (e.g. properties like 'being distinct from Boris Johnson'), as well as for its appeal to concreteness — which entails, according to Cowling, that abstract entities cannot have haecceities.

Furthermore, Hoffmann-Kolss (2019, p. 1000) remarks that according to Rosenkrantz's definition, some disjunctive haecceities are qualitative properties. One example is the property 'being identical to Joe Biden or being identical to Boris Johnson'. Hoffmann-Kolss argues that this property has neither Joe Biden nor Boris Johnson as concrete constituents, since its instantiation doesn't entail neither the existence of Joe Biden nor the existence of Boris Johnson. However, the property 'being identical to Joe Biden or being identical to Boris Johnson' is a paradigmatic case of non-qualitative property, and therefore Rosenkrantz's account is forlorn.

5.2.3 Loux's Definition

A third definition of qualitative properties can be found in Loux & Loux (1978, p. 133). According to Loux, a property P is non-qualitative if and only if there is some relation R and some individual x such that, necessarily, for any possible individual y: y instantiates P if and only if y bears R to x.[2]

Although Loux's definition delivers the right verdict on numerous prop-

[2] This definition is similar to Adams's (1979, 7–8) second definition of qualitative properties. (See footnote 1.)

erties, Hoffmann-Kolss (2019, p. 1002) argues that it wrongly identifies any qualitative properties which happen to be cointensional with some non-qualitative property as non-qualitative. One of the many examples offered by Hoffmann-Kolss (2019) is that of the properties 'being red' and 'having the same color which the Chinese flag actually has'.

It is easy to see that these two properties are cointensional, namely: they have the same extension at every possible world. This is because it is true in every possible world that red is the color which the Chinese flag actually has. (Or equivalently: it is a necessary truth that the Chinese flag is actually red.) And therefore, every possible object is such that it is red if and only if it has the same color which the Chinese flag actually has. However, while 'being red' is a paradigmatic case of qualitative property, 'having the same color which the Chinese flag actually has' seems to be non-qualitative.

And given this, it is true that there is some relation R, namely: 'having the same color which — actually has', and some individual x, namely: the Chinese flag, such that for any possible individual y, y instantiates 'being red' if and only if y bears the relation 'having the same color which — actually has' to the Chinese flag. Therefore, according to Loux's definition, the property 'being red' is non-qualitative.

It is important to notice that this is not just a challenge to Loux's view. More importantly, it is an argument to the effect that the distinction between qualitative and non-qualitative properties is inherently hyperintensional, where "[a]n X/Y distinction [among properties] is hyperintensional iff there are cointensional properties P and Q, such that P is an X-property and Q is a Y-property". (Hoffmann-Kolss 2015, p. 337) In support of this conclusion, Hoffmann-Kolss (2019) offers numerous examples of tuples of cointensional (i.e. necessarily coextensional) properties falling on different sides of the divide. One example mentions *qualitative essences*, i.e. combinations of qualitative properties which can be instantiated by only one individual, and which that individual instantiates in all possible worlds in which it exists. Hoffmann-Kolss's (2019) argument from qualitative essences to individuating qualitative and non-qualitative properties hyperintensionally is as follows. Suppose that individual essences exist, and suppose further that Joe Biden, say, has an individual essence: E. Consider now the haecceitistic property 'being identical to Joe Biden'. Clearly, this property is cointensional with E, for any world is such that 'being identical to Joe Biden' is instantiated if and only if E is instantiated, and there is only one object who could instantiate any of these properties: Joe Biden. However, while 'being identical to Joe Biden' is paradigmatically non-qualitative, E is by definition qualitative, and therefore they are distinct properties. It follows that an intensional individuation of properties is too coarse-grained to capture the distinction between qualitative and non-qualitative properties.

Another example given by Hoffmann-Kolss (2019) relies on the existence of some properties which are qualitative and cannot be instantiated by individuals only contingently. According to Hoffmann-Kolss, one such property is the property 'being an electron'. If this property could have been instantiated only contingently, then it would have been possible that there was an electron which could have been something else, like for example a proton. However, this is impossible for Hoffmann-Kolss. Notice now that the property 'being an electron', which is qualitative, is cointensional with the (infinitely complex) disjunctive property 'being e_1 or being e_2 or ...', where e_1, e_2, ... are all and only the possible individuals which are electrons. But again, this last property is non-qualitative.

I take this to show, with Hoffmann-Kolss (2019, p. 998), that any successful account of qualitative and non-qualitative properties should be sensitive to the fact that the distinction between these properties is hyperintensional. This is an important point, to which I will return in Section 5.3.3, where I will argue that the account I am proposing is not just a rehearsal of an old suggestion made, parenthetically, in Fine (1977).

5.2.4 Khamara's Defitinion

One last definition comes from Khamara (1988, p. 145), according to which "[a] property P is non-qualitative if and only if there is at least one individual, y, such that, for any individual, x, x's having P consists in having a certain relation to y".[3]

Khamara's approach has been challenged by Hoffmann-Kolss (2019, p. 1003) on the basis of its dependence on possibly inconsistent intuitions about the *consist-in* relation. Hoffmann-Kolss rightly remarks that intuitions about what a property consists in may vary. One might have the intuition that the property 'being Joe Biden' consists in standing in the identity relation to Joe Biden, and at the same time have the intuition that instantiating Joe Biden's qualitative essence also consists in standing in the identity relation to Joe Biden. In this case, Khamara's definition would classify Joe Biden's

[3]Loux's and Khamara's accounts see non-qualitative properties as relational. Another relational account of non-qualitative properties can be found in Rodriguez-Pereyra (2006, p. 205).

qualitative essence, which is by definition a qualitative property, as non-qualitative.

5.3 A New Idea

Cowling (2015) has recently argued that the failure of virtually all reductive definitions of qualitative properties proposed so far constitutes a reason for embracing a non-reductive view of qualitative and non-qualitative properties where haecceities are both ontologically and conceptually primitive.[4]

Although interesting, Cowling's solution is likely to raise some eyebrows. For haecceities do not have many friends, and surely an ontology without a primitive qualitative distinction should be preferred, for reasons of ideological economy, over one where, *ceteris paribus*, the distinction is not further reducible.[5]

In this Section I propose a new account of qualitative and non-qualitative properties which, I hope, can be used by the reductionist within a background theory where the distinction between qualitative and non-qualitative properties is not conceptually primitive. I suggest that we can define non-qualitative properties as those properties which stand constant against identity variations. More precisely, I will argue that a property P is non-qualitative whenever, given some individual x, whether x has P will vary when considering worlds/(im)possibilities which only disagree with respect to the identities

[4]Cowling is not the only author endorsing some form of primitivism when it comes to qualitative and non-qualitative properties. Another author which has argued in favor of a primitivist account of the distinction between qualitative and non-qualitative properties is Diekemper (2009).

[5]Arguments against haecceities can be found, *inter alia*, in: Fine (1985b), Lowe (2003), and Moreland (2013).

they assign to the relevant individuals.[6] In what follows, I will talk about this feature as a form of 'invariance under identity assignments'.

It is important to remark at the outset that by 'identity' I don't intend to refer to the *relation* of identity. Instead, I mean to refer to what Locke (1689, III.3.15) would paraphrase as "the very being of any thing, whereby it is, what it is".[7]

5.3.1 Invariance under Identity Assignments

The main idea I want to suggest is that a property P is qualitative if and only if it is invariant under any identity assignment. To a first approximation, we can consider an identity assignment as a function $i : D \to H$ from a collection of entities D to a collection of identities H. What identities are is an interesting philosophical question, but one which is orthogonal to the present purposes. However, as a useful analogy, one might think of identities as markers. Suppose you have a collection D of unmarked entities and a collection H of markers: to a first approximation, an identity assignment is just a way to assign markers in H to entities in D.

Now, the thought that if it is impossible that Joe Biden could have been Boris Johnson then there is an impossible world w such that the entity which is Joe Biden at the actual world is Boris Johnson at w, gives us an additional piece of information about identities and their relations to individuals. Namely, it tells us that, given an individual d, d's identity will depend also

[6] In this Chapter, I assume for simplicity that possible worlds are possibilities and impossible worlds are impossibilities. Nothing will hinge on this assumption.

[7] The distinction between these two senses of 'identity' is discussed, among others, in Lowe (2016, p. 52–53). On Locke's understanding of the notion of 'essence', see: Owen (1991).

on the world in which claims about d are being considered. Therefore, to a second approximation, we say that an identity assignment is a function $i : D \times (W \cup I) \to H$, where: D is a collection of individuals, W and I collections of possible and impossible worlds respectively, and H a collection of identities. With this second approximation, we can make sense of the idea that the identity of a given individual might shift across worlds. In particular, since we believe that identities are necessary, we say that for any individual d, any possible world w, and any identity h: if h is d's identity in w then for any possible world v, h is d's identity in v. Since this doesn't hold for impossible worlds, we can make sense of the idea that in these worlds the identities of our individuals might shift.

To better appreciate how invariance under identity assignment is related to qualitative properties, consider the following example. Take the property 'being Boris Johnson's father', which at the time of writing is instantiated by Stanley Johnson. Suppose that, *per impossible*, Boris Johnson was not Boris Johnson, but Joe Biden. Then, Stanley Johnson would not instantiate 'being Boris Johnson's father'. Instead, Stanley Johnson would instantiate the property 'being Joe Biden's father'. The property 'being Boris Johnson's father' is therefore non-qualitative, for the fact that it is instantiated by Stanley Johnson depends on the fact that Boris Johnson is indeed Boris Johnson, and not Joe Biden. That is: were Boris Johnson to be Joe Biden, the extension of the 'being Boris Johnson's father' would change.

In the above example, Joe Biden would instantiate many of the properties which Boris Johnson instantiates, among which: the property 'having been the Prime Minister of the United Kingdom', the property 'having never been the President of the United States of America', and the property 'having attended Eton College'. The idea behind the example, and my account of qualitative properties, is not the exchange of *entities*: we are not swapping, as it were, a baby Joe Biden with a baby Boris Johnson, and let them grow in a different environment. The idea is the exchange of their *identities*, like in an identity theft that has gone all too well: as if, after counterfeiting Boris Johnson's ID card, Joe Biden *did indeed become* Boring Johnson, in a sense of 'become' that includes the property 'always having been Boris Johnson'.

5.3.2 Identity Assignments and Haecceitism

My definition of non-qualitative properties is easily translatable in the language of Haecceitism, which is the thesis that there are maximal possibilities which include all the same qualitative possibilities, and yet differ with respect to the non-qualitative possibilities they include. (See Cowling 2017.) Then, my suggestion is equivalent to the thesis that a property P is non-qualitative if and only if its extension varies across possibilities and impossibilities which differ haecceitistically. (Possibilities and impossibilities differ haecceitistically only when they differ at most with respect to *which* entity is *which*, among the entities they represent as being one way or another).

This, I hold, is evidence that my definition of non-qualitative properties is promising. On the one hand, in fact, its translation in the language of Haecceitism makes it extremely simple and intuitive. On the other hand, the translation highlights the fact that my definition is in line with one of the most common intuitions regarding non-qualitative properties, namely: that unlike qualitative properties, non-qualitative properties are able, at least in principle, to distinguish between qualitative duplicates.[8] (Notice that by definition, if (im)possibilities P_1 and P_2 differ haecceitistically, then P_1 and P_2 are qualitative duplicates.)

One last remark about the connection between my account of non-qualitative properties and Haecceitism. As it stands, Haecceitism is a thesis that only few authors explicitly endorse.[9] Also, Haecceitism comes with different degrees of commitment, so to say. Take for instance Black's (1952) symmetrical world w, discussed in Sections 1.1.1, containing only two indiscernible iron spheres two miles away from each other. It seems rather uncontroversial that if w represents a genuine possibility, call it P_1, then there is another possibility P_2 such that P_1 and P_2 disagree only with respect to *which* sphere is Castor and *which* one is Pollux.[10] More difficult to accept is the existence of a possibility P_3 such that the only difference between the possibility represented by the actual world and P_3 is in the fact that according to P_3 Joe Biden never existed and I have all the qualitative properties that Joe Biden has according to the possibility represented by the actual world. (Among other things, P_3 would violate the Principle of the Necessity of Origin.) Even more controversial is to accept the existence of a possibility P_4 such that the only difference between the possibility represented by the actual world and P_4 is in the fact that according to P_4 Joe Biden is a unicellular organism, or an electron.

[8]See, among others: Eddon (2010, p. 317), and Simmons (2020, p. 3066).

[9]For arguments against Haecceitism, see: Armstrong (1989), Dasgupta (2009), Forbes (1985), and D. Robinson (1989).

[10]See Chapter 4 for an argument in support of the existence of such possibility P_2.

This is to say: the great majority of authors would argue that, given a possibility P, most of the possibilities which haecceitistically differ from P are indeed impossibilities. My definition relies on such (im)possibilities, as well as on (im)possibilities which represent other (im)possibilities as being actualised. (This is because I take that possibilities and impossibilities are included among the individuals a non-qualitative property might depend on.)

This might seem, at a first glance, a downside of my definition. However, I don't think it is. For although it is still an open question whether properties and relations should be individuated hyperintensionally, we have seen in Section 5.2.4 that Hoffmann-Kolss (2019) makes a good case for the existence of qualitative properties which are cointensional with non-qualitative ones — and a good case, as a consequence, for the necessity of a hyperintensional account of qualitative and non-qualitative properties. If Hoffmann-Kolss is right, then one should prefer my account, which classifies properties on the basis of how their extensions are represented by possibilities and impossibilities alike, over any account which identifies cointensional properties.

5.3.3 Fine's Suggestion

A suggestion which looks similar to my definition comes from Fine (1977, p. 174). Fine's account too ties the non-qualitative character of properties to their dependence on individuals' identities, and makes the notion of dependence precise by using the language of *possible worlds* and *automorphisms*. In Fine's (1977, p. 136) vocabulary, an automorphism is defined as a function between possible worlds which assigns any world w to a world v such that v

and w agree with respect to everything except, perhaps, with respect to the *identity* of the individuals they contain. That is: v and w contain the same individuals, and yet disagree about which individual is which.

In particular, Fine suggests that a property P such that at some possible world w it is the case that some individual x has P depends on the identity of some individual(s), and is therefore non-qualitative, if and only if its extension changes across possible worlds which are related to w by some automorphism. More formally: a property P such that at some possible world w it is the case that some individual x has P depends on the identity of some individual(s), and is therefore non-qualitative, if and only if there is some possible world v and some automorphism f such that $v = f(w)$ and the extension of P at w is different from the extension of P at v.

To see how Fine's account works, consider again the property 'being Boris Johnson's father'. Call this property 'P', and let w be a world where there is some individual x which has P. Now consider a possible world v which agrees with w with respect to everything except for the fact that the entity that is Boris Johnson in w is Joe Biden in v. Then, there is an automorphism f such that $v = f(w)$. By construction, v is such that x doesn't have P. Therefore, since there is an automorphism f and a possible world v such that $v = f(w)$ and the extension of P at w is different from the extension of P at v, we conclude that P is non-qualitative. This is in line with the common treatment of properties like 'being Boris Johnson's father'.

Fine's account of qualitative properties and my definition are indeed similar. However, they deliver different results: there are many properties, in fact, which Fine's account misclassifies as qualitative while my account correctly classifies as non-qualitative, as I explain below.

For example, according to Fine's account, any impossible property (by which I mean any property which is not possibly instantiated) is qualitative. Take, for example, the property 'being a talking even number'. Clearly, there is no possible world which contains a talking even number. (That is: for any possible world w, the extension of 'being a talking even number' is, according to w, the empty set.) Therefore, given a possible world w and a class F of automorphisms on w, it is easy to see that for all possible worlds v such that $v = f(w)$ for some f in F, the extension of 'being a talking even number' doesn't change. Therefore, according to Fine's definition, 'being a talking even number' is a qualitative property. We will see that, on this property, my account will deliver the opposite result. On the assumption that the property 'being even' is inherently relational, and in particular is the property 'being divisible by 2 without reminder', my account characterises the property 'being even' as non-qualitative. And since the property 'being a talking even number' is a conjunctive property where one of the conjuncts is a non-qualitative property, then 'being a talking even number' is non-qualitative too.

Also, suppose that you believe, with Karofsky (2021), that everything is necessary. Then, you believe that the entire logical space is exhausted by the actual world, that is: that the actual world is the only possible world. If this is the case then the only automorphism you can accept is the trivial automorphism $f(w) = w$, where w is the actual world. And as a consequence Fine's criterion delivers that any property which is instantiated in the actual world, like the property 'being identical to Boris Johnson', is qualitative.

Finally, the main shortcoming of Fine's account seems to be that it identifies as qualitative some properties which are intuitively non-qualitative, like the property 'having the same color which the Chinese flag actually has', and which happen to be cointensional with some qualitative property.[11] Again, the property 'having the same color which the Chinese flag actually has' and the property 'being red' are cointensional, which means that they have the same extension at every possible world. However, 'being red' is a qualitative property, and 'having the same color which the Chinese flag actually has' is not. And since it is a necessary truth that the Chinese flag is actually red, and this truth doesn't depend on the identity of the Chinese flag, then the extension of the property 'having the same color which the Chinese flag actually has' will be invariant under all automorphisms.

To see this, let w be the actual world, and f an arbitrary automorphism such that $f(w) = v$ for some possible world v which agrees with w about everything except the identity of the Chinese flag. No matter *which* entity is the Chinese flag in v, and no matter what color it is, it will be true in v that the *actual* Chinese flag is red. Therefore, the extension of the property 'having the same color which the Chinese flag actually has' in v will be different from the extension of the property 'having the same color which the Chinese flag actually has' in w if and only if the extension of the property 'being red' in v is different from the extension of the property 'being red' in w. Which means that either 'having the same color which the Chinese flag actually has' and 'being red' are both qualitative, or they are

[11] We have rehearsed this challenge, put forward by Hoffmann-Kolss (2019, p. 1002), in Section 5.2.3.

both non-qualitative. And since the extension of the property ‘being red’ doesn’t change, then also the extension of the property ‘having the same color which the Chinese flag actually has’ doesn’t change. Therefore, ‘having the same color which the Chinese flag actually has’ is identified by Fine as a qualitative property.

By introducing impossible worlds and considering possible and impossible worlds among the entities which can be subject to exchanges of identities, my account solves all these issues and agrees with Hoffmann-Kolss’s intuitions about this and other pairs of cointensional properties. It does so because the notion of identity assignment is not restricted to worldbound individuals, and it is defined for every collection of individuals whatsoever. (Where an individual is, following French and Krause (2006, p. 248), any entity to which identity applies.) By considering alternative identity assignments on *any* collections of individuals, we can not only exchange the identities of worldbound individuals, but also the identities of possible and impossible worlds. (In this sense, my account can be thought of as a generalisation of Fine’s 1977 suggestion.) Notice in fact that according to French and Krause’s definition of individuals, both possible and impossible worlds, as well as numbers, sets, events, and moments count as individuals. This generalisation will allow us to analyse modal properties, for example, as properties involving quantification over possible and impossible worlds.

5.4 The Logic of Qualitative Properties

In this Section I provide the start of a formal framework capable of distinguishing between qualitative and non-qualitative properties in line with the account I have proposed in this Chapter. Even disregarding the issue of qualitative and non-qualitative properties, this framework is interesting in itself, given than it can be applied to any debates in which a prominent role is played by entities which, in some way or other, depend on the identity of individuals. Since for my purposes an intuitive understanding of the framework will suffice, I will not aim at comprehensiveness, relegating questions about the logical properties of this system to future work.

The framework I define in this Section is inspired by Priest's (2008) Contingent Identity Modal Logic, but differs from it in important respects. (I will expand on this is Section 5.4.3.) Furthermore, it is also compatible with primitivist accounts of the qualitative properties, like the one put forward in Cowling (2015). Indeed, as we will see, my framework provides a theoretically fruitful way to tell, given a property P, whether P is qualitative or not, and this is a feature that all the extant primitivist approaches, Cowling's included, currently lack.

5.4.1 Transparent and Luminous Regimentations

When we take a formula '$\varphi(x)$' in a language $\mathcal{L}$ to express a certain property P, we say that '$\varphi(x)$' is a *regimentation* of P in $\mathcal{L}$. In a sufficiently expressive language $\mathcal{L}$, many $\mathcal{L}$-formulas can regiment an expressible property P. By way of example, say that $\mathcal{L}$ is a common first-order language and P the property 'being Donald Trump's father'. The formula '$F(x)$', where 'F' is a unary predicate, is a possible regimentation of P in $\mathcal{L}$.

This regimentation, however, inevitably hides the similarity between P and the property Q of 'being Hillary Clinton's father'. This is because, no matter how we regiment Q, there is nothing in the formula '$F(x)$' from which we can deduce that P and Q share a relation, e.g. 'being the father of —'. However, were we to regiment P and Q as '$R(x, d)$' and '$R(x, h)$' respectively, where 'd' is a constant denoting Donald Trump and 'h' a constant denoting Hillary Clinton, their similarity would be accounted for by the similarity of their regimentations.

In this sense, we can say that some regimentations are more *transparent* than others, in the sense that they better mirror the nature of the property we want to regiment. Given an expressive enough language $\mathcal{L}$, we say that a regimentation '$\varphi(x)$' of some property P in $\mathcal{L}$ is *luminous* when it exactly matches the structure of P.[12]

5.4.2 Language & Interpretations

With this in mind, let $\mathcal{L}$ be a standard first-order language with infinitely many individual constants, variables, relation symbols, and the usual logical constants plus identity. Well-formed formulas in $\mathcal{L}$ are defined as usual. We define an interpretation to be a sixtuple $\mathcal{I} = \langle D, H, W, I, i, v \rangle$, where: D is a non-empty domain of quantification, H is a non-empty set of entities

[12]Clearly, given some property P and a regimentation '$\varphi(x)$' of P, whether one will consider '$\varphi(x)$' a luminous regimentation of P will depend on what the ontology they endorse has to say about P.

which we use to represent the identities of the individuals in D, W is a non-empty set of possible worlds, and I is a non-empty set of 'compositional' impossible worlds, i.e. impossible worlds where the truth of complex formulas is compositional and where, as we will see in a moment, the clauses for determining the truth-value of a complex formula given the truth-values of its immediate subformulas are the same as in standard possible worlds.[13] Finally, $i : D \times (W \cup I) \to H$ is a function from individuals and worlds to identities, assigning each individual an identity at each world, and v is our interpretation function.

We add some constraints on our interpretations. First, we stipulate that any interpretation $\mathcal{I}$ is such that, for all $d \in D$: $i(d, w) = i(d, v)$ for all $w, v \in W$. (This means that all possible worlds will agree with respect to the identity of all the individuals in D.) Second, since we want to quantify over possible and impossible worlds in our language, we say that $(W \cup I) \subseteq D$. Third, we say that for every individual constant c in $\mathcal{L}$, $v(c) \in D$, and for all $d \in D$ there is a constant c_d in $\mathcal{L}$ such that $v(c_d) = d$. Then, we say that for every world $w \in W \cup I$, and for every $n-$place predicate P^n, the interpretation $v_w(P^n)$ of P^n at w is a pair $\langle E^+_w(P^n), E^-_w(P^n)\rangle$, where: $E^+_w(P^n)$ is the extension of P^n at w and $E^-_w(P^n)$ the anti-extension of P^n at w.[14]

[13] As we will see shortly, 'compositional' impossible worlds are all we need to specify a framework capable of distinguishing between qualitative and non-qualitative properties. Nonetheless, it would be interesting to expand the present framework so that it contains 'non-compositional' impossible worlds. For more on compositionality and impossible worlds, see: Berto & Jago (2013, §4.3 and §8.5).

[14] My frameworks combines ideas from Priest's Logic of Paradox (see, among others, Priest 1997), and Priest's Contingent Identity Modal Logic (see Priest 2008, ch. 17). This is because I want impossible worlds to make contradictions true, and tautologies false. And one way to do this is by specifying the interpretation of predicates in terms of their extensions and anti-extensions.

Furthermore, for every world $w \in W \cup I$, and for every $n-$place predicate P^n, $\upsilon_w(P^n) \in \wp(H^n) \times \wp(H^n)$ — that is: the interpretation of P^n at w is a member of the product of the powerset of H^n.[15] We say that for every possible world w, and for every $n-$place predicate P^n: $E^+_w(P^n) \cap E^-_w(P^n) = \emptyset$, and $E^+_w(P^n) \cup E^-_w(P^n) = H^n$. None of these conditions apply to impossible worlds. Finally, we say that for all possible worlds $w \in W$, $E^+_w(=)$ is the set $\{\langle h, h\rangle : h \in H\}$. For impossible worlds, the extensions and anti-extensions of the identity symbol '=' are unconstrained.

Given an interpretation $\mathcal{I} = \langle D, H, W, I, i, \upsilon\rangle$, we assign to every closed formula φ in $\mathcal{L}$ a semantic value $V_{\mathcal{I},w}(\varphi)$ in the set $\{\{1\}, \{1, 0\}, \{0\}, \emptyset\}$[16] for every possible or impossible world w as follows. If φ is of the form $Pc_1, ..., c_n$ and $\mathcal{I}$ is an interpretation and w a possible or impossible world, then:

$1 \in V_{\mathcal{I},w}(\varphi)$ iff $\langle i(\upsilon(c_1), w), ..., i(\upsilon(c_n), w)\rangle \in E^+_w(P)$, and

[15] At this point one might object that, contrary to common sense, predicates in $\mathcal{L}$ are true of identities, and not of individuals in the domain, that is: that although the formulas in $\mathcal{L}$ quantify over entities in D, what they predicate of, describe, and speak of, are entities in H — and this is a shortcoming of the present account. I am happy to concede the point. However, the present framework is purely instrumental: whether it successfully distinguishes between qualitative and non-qualitative properties is "all that matters" for my purpuses. Second, there is something to the idea that identities play a role in whether predicates apply to entities in our domain — and I take this to be the basic intuition behind my account of qualitative properties. What role identities actually play is an open question, and one whose answer requires a more comprehensive metaphysics of identities. Still, the basic intuition here is that the property 'being David Lewis' is true of David Lewis insofar as David Lewis *is* (in the sense of identities) David Lewis. Third, notice that we can associate to each predicate's extension and anti-extension (in the sense defined above), a collection of objects from the domain, respectively. For instance, for any monadic predicate P and any world w there will be a set P^+_w such that $P^+_w = \{x \in D : i(x, w) \in E^+_w(P)\}$ and a set P^-_w such that $P^-_w = \{x \in D : i(x, w) \in E^-_w(P)\}$. These are sets of individuals in D, and one can regard these sets as the *true* extension and anti-extension of every predicate in $\mathcal{L}$. (It should be easy to see how the above generalises to n-place predicates, for any n.) This should dispel the worries completely.

[16] The addition of the empty-set to the usual truth-values of LP logic is motivated by the fact that we want some impossible worlds to be such that the extension and anti-extension of some given predicate P^n are not jointly exhaustive. This will prove useful in showing that some tautological properties are non-qualitative.

$0 \in V_{\mathcal{I},w}(\varphi)$ iff $\langle i(\upsilon(c_1), w), ..., i(\upsilon(c_n), w)\rangle \in E_w^-(P)$.

The extension of this notion to formulas containing connectives and quantifiers is as follows. For all formulas φ and ψ in $\mathcal{L}$:

$1 \in V_{\mathcal{I},w}(\neg\varphi)$ iff $0 \in V_{\mathcal{I},w}(\varphi)$;

$0 \in V_{\mathcal{I},w}(\neg\varphi)$ iff $1 \in V_{\mathcal{I},w}(\varphi)$;

$1 \in V_{\mathcal{I},w}(\varphi \wedge \psi)$ iff $1 \in V_{\mathcal{I},w}(\varphi)$ and $1 \in V_{\mathcal{I},w}(\psi)$;

$0 \in V_{\mathcal{I},w}(\varphi \wedge \psi)$ iff $0 \in V_{\mathcal{I},w}(\varphi)$ or $0 \in V_{\mathcal{I},w}(\psi)$;

$1 \in V_{\mathcal{I},w}(\forall x\varphi(x))$ iff for all $d \in D$: $1 \in V_{\mathcal{I},w}(\varphi[c_d/x])$, where '$\varphi[c_d/x]$' is the formula obtained by uniformly replacing all the free occurrences of 'x' in $\varphi(x)$ with 'c_d';

$0 \in V_{\mathcal{I},w}(\forall x\varphi(x))$ iff for some $d \in D$: $0 \in V_{\mathcal{I},w}(\varphi[c_d/x])$.

Disjuction and material conditional, as well as the existential quantifier, are defined as usual in terms of conjunction, negation, and the universal quantifier. Finally, we say that an interpretation $\mathcal{I}$ at a world w satisfies φ (write, '$\mathcal{I}, w \vDash \varphi$'), if and only if $1 \in V_{\mathcal{I},w}(\varphi)$. Though we won't need it below, we can define a natural relation of entailment as follows: a set of (closed) formulas Σ entails a (closed) formula φ whenever for every interpretation $\mathcal{I}$ and possible world $w \in W$: $\mathcal{I}, w \vDash \varphi$ if $\mathcal{I}, w \vDash \sigma$ for all $\sigma \in \Sigma$.

5.4.3 Contingent Identity Modal Logic

As mentioned above (Section 5.4), the formal framework set up in Section 5.4.2 is inspired by Priest's (2008) framework for Contingent Identity Modal

Logic. However, there are important differences between the two, which is important to spell out before going further.

Unlike Priest's (2008) modal language, the language defined in Section 5.4.2 doesn't have logical constants for possibility and necessity. $\mathcal{L}$ is in fact a non-modal language, tailored to a conception of properties according to which modal properties involve quantification over possible and impossible worlds.[17] This is also the main reason why we don't have an accessibility relation between worlds in our interpretations (while Priest does). Since we just want to express properties, and every property will be luminously regimented without modal symbols, we don't need an accessibility relation in our interpretations. However, this doesn't mean we cannot have one. There is no reason I can see why $\mathcal{L}$ couldn't be expanded to include symbols for modalities, and a relation between possible (and impossible?) worlds added to our current interpretation.

Somewhat relatedly, since we want to be able to quantify over possible and impossible worlds, we have them in our domains as individuals which have their own identities. This is part of the reason why we don't define, as Priest (2008, p. 368) does, the individuals is D as functions from worlds to identities. For if we wanted the elements of our domain to be functions from worlds to identities and at the same time we wanted to have worlds

[17]Note that this is because I believe modal properties are properties which quantify over possible worlds. If you believe this is not the case, and you believe that no property quantifies over worlds, then you could work with interpretations where $W \cap D$ is empty. This will probably give different results about the status of (at least) modal properties. But this has to be expected, and it is in no way a shortcoming of the present account: for as I remarked above, whether one will believe a property is qualitative or non-qualitative will heavily depend on their intuitions about what is the nature of the property in question. This, in turn, will bear consequences for which regimentation of the relevant property one recognises as luminous.

in our domain, we would end up with functions which are not well-founded. This is of course not an issue *per se*, since we could have a non-well-founded set theory as our ambient theory of sets. However, in order to keep things simple, we stick with ZFC for the moment and define a function i for each model assigning an identity to each individual in D relative to each world. A second reasons why I decided to go for a further function instead of treating the individuals in D as being themselves functions is that the elements of D are meant to represent real-world individuals, and many of them (the sun, an hydrogen atom, a galaxy), are not functions.

Finally, since we believe that identity is necessary, we introduce the constraint that any two possible worlds always assign the same identity to any individual in the domain. This is not the case in Priest's framework, for as the reader can infer from its name, Priest's logic is one which is meant to capture the idea that identity is contingent. And since we want to accommodate for exchanges in identities and we believe identity is necessary, we have introduced impossible worlds into the interpretations. Unlike possible worlds, we let impossible worlds diverge in their identity assignment.

5.4.4 Duplicate Worlds

With this in mind, we define the notion of 'duplicate worlds'. Informally, given an interpretation $\mathcal{I}$: possible or impossible world w and v are φ-duplicates in $\mathcal{I}$ whenever they agree with respect to everything except, at most, with respect to the identities they assign to the elements of the domain referred to by the constants in φ. We define duplicate worlds as follows:

Duplicate Worlds (Definition): Given an interpretation $\mathcal{I} = \langle D, H, W, I, i, v \rangle$, a formula φ in $\mathcal{L}$, and worlds w and v in $W \cup I$, we say that v is a φ-duplicate of w if and only if: (1) for all predicates P^n in φ, $n \in \mathbb{N} : v_w(P^n) = v_v(P^n)$, and (2) for all $d \in D$, if there is no constant c in φ such that $v(c) = d$, then $i(d, w) = i(d, v)$.

Given an interpretation $\mathcal{I}$ and a formula φ in $\mathcal{L}$, we indicate that worlds w and v in $W \cup I$ are φ-duplicate by writing '$w \sim_\varphi v$'.[18] It should be easy to see that we can use duplicate worlds to capture changes in the identities of the relevant individuals.

5.4.5 The Definition of Qualitative Properties

We can then present a technical definition of qualitative properties:

Qualitative Properties (Definition): Let P be a property and $\varphi(x)$ a luminous regimentation of P in $\mathcal{L}$. P is qualitative if and only if, for all interpretations $\mathcal{I} = \langle D, H, W, I, i, v \rangle$ and all worlds $w \in W \cup I$: if $\mathcal{I}, w \vDash \varphi[c/x]$, then $\mathcal{I}, v \vDash \varphi[c/x]$ for all $v \sim_{\varphi(x)} w$, where '$\varphi[c/x]$' is the formula obtained by uniformly replacing all the free occurrences of 'x' in $\varphi(x)$ with 'c', and the constant 'c' in $\varphi[c/x]$ was not already in $\varphi(x)$.

Finally, we say that a property is non-qualitative whenever it is not qualitative.

[18] So defined, 'duplication' is an equivalence relation. Hence, given a formula φ in $\mathcal{L}$, and an interpretation $\mathcal{I}$ such that $\mathcal{I}, w \vDash \varphi$ for some w in $W \cup I$, the class of worlds v such that $v \sim_\varphi w$ is an equivalence class under the relation of φ-duplication.

5.5 A Simple Application

In this Section, I show an application of the above framework to the property 'being Boris Johnson's father', which I mentioned in Section 5.3.1. The aim is to highlight how the reasoning we employed in Section 5.3.1 to reach the conclusion that 'being Boris Johnson's father' is a non-qualitative property is mirrored by the formal framework presented in Section 5.4.

So consider again the property P of 'being Boris Johnson's father'. P is a relational property: in particular, it is the property of bearing the relation 'being the father of —' to whatever happens to be Boris Johnson. If this is correct, then P can be luminously regimented in $\mathcal{L}$ as the formula: '$F(x, b)$'. To see that P is non-qualitative, consider an interpretation $\mathcal{I} = \langle D, H, W, I, i, v \rangle$ with worlds $w, v \in W \cup I$ such that:

- $D = \{d_1, d_2\}$;
- $v(a) = d_1$ and $v(b) = d_2$;
- $E^+_w(F) = E^+_v(F) = \{\langle h_1, h_2 \rangle\}$;
- $E^-_w(F) = E^-_v(F)$;
- $i(d_1, w) = i(d_1, v) = h_1$;
- $i(d_2, w) = h_2$ and $i(d_2, v) = h_3$;

Since $\langle i(d_1, w), i(d_2, w) \rangle \in E^+_w(F)$, then it is the case that $\mathcal{I}, w \vDash Fab$. However, since $\langle i(d_1, v), i(d_2, v) \rangle \notin E^+_v(F)$, then it is *not* the case that $\mathcal{I}, v \vDash Fab$.

Since clearly, in $\mathcal{I}$, $v \sim_{Fxb} w$, then by the definition of qualitative properties in Section 5.4.5, P is non-qualitative.

In Section 5.3.3 I argued that the account of qualitative and non-qualitative properties I have presented in this Chapter is to be preferred to Fine's view, in that (1) it doesn't uniformly classify impossible properties as qualitative, (2) it is general enough to overcome Hoffmann-Kolss's (2019) challenge of cointensional qualitative and non-qualitative properties, and (3) it is inherently hyperintensional. With the formal machinery defined in Section 5.4, we can see this more clearly.

Recall that Fine (1977) defines automorphisms as functions between possible worlds: given a possible world w and an automorphism f on w, $f(w)$ is a world which agrees with w with respect to everything except perhaps with respect to the identities of the individuals it contains. This means that Fine's automorphisms allow us only to look at situations where world-bound individuals, but not worlds themselves, exchange identities. With the interpretation of identity assignments our formal framework affords, this limitation is removed. This can help us analyse modal properties, like 'being an actual donkey', as properties that involve quantification over possible worlds. The idea is simple. Suppose you believe that the property 'being an actual donkey' is a property that relates a possible individual and a possible world. I, for example, understand this property as the property 'being a donkey according to the actual world'. So for me, a luminous regimentation of the property 'being an actual donkey' in $\mathcal{L}$ would look like the formula: '$R(x, D, @)$', where 'R' is a relational constant denoting the relation of 'according to —', 'D' a unary predicate denoting the property 'being a donkey', and '@' a constant denoting the actual world. Under this understanding of the property, it comes natural to ask whether its extension changes when we exchange the identities of the individuals involved. And since the only individual involved is the actual world, it makes sense to ask if the property would have a different extension were the actual world a different individual. This, which seems to me impossible in Fine's account, is instead possible in the framework presented in Section 5.4.

5.6 Philosophical Remarks

I want to conclude this Chapter with two general remarks. The first has to do with complex properties, like the property 'being five miles away from Joe Biden *and* being tall'. The second has to do with some peculiar kinds of properties, which, following Cowling (2015), any good account of non-qualitative properties should be able to deliver a reasonable verdict about. In this Section, I first show that my new account characterises complex properties in a way which is in line with the main intuitions in the literature. Then, I test it against the list of candidate kinds of non-qualitative properties suggested in Cowling (2015, p. 283–286). I argue that my view correctly identifies as non-qualitative all those properties with respect to which the literature agrees. Furthermore, I argue that my suggestion is capable of providing new valuable intuitions when it comes to those properties whose status is still a matter of debate.

5.6.1 Complex Properties

It is easy to see that according to my account, tautological and contradictory properties do not come out as uniformly qualitative. This is because, since we are not restricting ourselves to the consideration of possibilities, we can define suitable models in which contradictory properties apply to some entities, and tautological properties fail to apply to some entities. Furthermore, since we allow the exchange of the the identity of possibilities and possible worlds, we can define collections of duplicate models in which the extension of some contradictory and tautological properties change.

I hold that this is a good feature of my account. For consider the two tautological properties 'being either red or not red' and 'being either identical to Joe Biden or distinct from Joe Biden'. It is clear that the first property is a disjunction of qualitative properties ('being red' and 'being not red') and the second a disjunction of non-qualitative properties ('being identical to Joe Biden' and 'being distinct from Joe Biden'). Now, one intuition which is shared in the literature is that disjuctive properties are qualitative if and only if both their disjuncts are qualitative. (See, among others: Adams 1979, p. 8). Therefore, the property 'being either red or not red' should be qualitative while the property 'being either identical to Joe Biden or distinct from Joe Biden' should be non-qualitative. And they so are, according to my definition.

In general, with respect to complex properties, my account delivers the following results:

1. A property P is qualitative if and only if its negation is.

2. A conjunctive property (P and Q) is qualitative only when both P and Q are qualitative.

3. A disjunctive property (P or Q) is qualitative only when both P and Q are qualitative.

Again, this is in line with the main intuitions in the literature (See Adams 1979).

5.6.2 Haecceities

According to Cowling (2015), any account of qualitative and non-qualitative properties worthy of its name should identify certain paradigmatic properties as non-qualitative, and somehow match the intuitions in the literature about other more problematic ones.

For a start, Cowling suggests that any good definition of qualitative properties should identify *haecceities*, *impure properties*, *negative haecceities*, and *disjuntive haecceities* as non-qualitative. One example of *haecceity* is the property 'being identical to Boris Johnson'. An example of *impure properties* is the property 'standing next to Joe Biden'. A paradigmatic *negative haecceity* is the property 'being distinct from Hillary Clinton'. Finally, an instance of *disjuntive haecceity* is the property 'being either identical with Boris Johnson or identical with Theresa May'.

My account correctly classifies haecceities as non-qualitative: they are in fact the simplest properties whose distribution changes across duplicate worlds. From what we said in Section 5.6.1, it also follows that my account correctly identifies both negative haecceities and disjunctive haecceities as non-qualitative.

Before we pass on to the other candidate kinds of non-qualitative properties listed in Cowling (2015), I want to clarify that in the last paragraph I have been using Adams's (1979, p. 6) definition of haecceities, according to which an haecceity "[...] is the property of being identical with a certain particular individual — not the property that we all share, of being identical with some individual or other, but my property of being identical with me, your property of being identical with you, etc." Many authors distinguish between the property 'being Donald Trump' and the property 'being identical to Donald Trump'. (See, among others, Cowling 2021 and Lewis 1986.) Although I personally believe, with Plantinga (2003), that such distinction is misguided, it is important to notice that the account presented in Section 5.4, being purely formal, is neutral on whether, given an entity x, the properties 'being x' and 'being identical to x' are indeed distinct.

And, unlike Fine's suggestion, my account is also neutral with respect to the existence of qualitative essences. (Fine's account is usually taken to entail that there are no qualitative essences.) This is, I believe, another point in favor of my view, which is immune to a criticism which Hoffmann-Kolss (2019, p. 1000) suggested against Rosenkrantz's account of non-qualitative properties, and which can be equally applied to Fine's suggestion, namely: that any good account of non-qualitative properties should be neutral with respect to whether individuals have qualitative essences.

5.6.3 Impure Properties

For what concerns impure properties, it should be easy to see that they are also non-qualitative according to my view. We have seen in Sections 5.3.1 and 5.6 that my account classifies the impure property 'being Boris Johnson's father' as non-qualitative. Almost all other impure properties can be luminously regimented in $\mathcal{L}$ by formulas that structurally resemble the regimentation of 'being Boris Johnson's father', which, the reader will remember, was the formula '$F(x, b)$'.

For instance, the impure property 'standing next to Joe Biden' can be luminously regimented in $\mathcal{L}$ by some formula '$S(x, j)$'. Similarly, the property 'being five miles from the Chrysler Building' can be luminously regimented as '$D(x, c)$'. Non-binary impure properties, like the property 'standing between Donald Trump and Joe Biden', can be luminously regimented in $\mathcal{L}$ by formulas like '$S(x, d, j)$'. It is an easy exercise to check that, given a regimentation '$S(x, d, j)$' for the property 'standing between Donald Trump and Joe Biden' in $\mathcal{L}$, there is an interpretation $\mathcal{I} = \langle D, H, W, I, i, v\rangle$ and worlds w, v in $W \cup I$ such that $\mathcal{I}, w \vDash S(a, d, j)$ and $\mathcal{I}, v \nvDash S(a, d, j)$.

5.6.4 Tense and Modal Properties

What about *tense* and *modal* properties, like 'being ill on the 10th of July 2019' or 'being a possible donkey'? Cowling (2015, p. 283) argues that if we accept both eternalism and modal realism and we believe that there is a fundamental difference between actual entities and possible entities, as well as between present entities and past or future entities, we should classify tense and modal properties as non-qualitative.

My account respects this intuition. If we regiment tense properties as involving quantification over units of time and modal properties as involving quantification over possible worlds and/or possibilities, we can see that both tense and modal properties are identified as non-qualitative.

One way to luminously regiment the property 'being ill on the 10th of July 2019' in $\mathcal{L}$ is by means of the formula '$\exists y(y = 10/07/2019 \wedge I(x, y))$', while to regiment the property 'being possibly a donkey' we could use the formula '$\exists w(R(@, w) \wedge D(x, w))$'. (Clearly, these are simplified formulas, but they should nonetheless give an idea of how $\mathcal{L}$ captures tense and modal properties.)

5.6.5 Structural Properties

After tense and modal properties, Cowling considers *structural properties*, like 'being distinct from something' and 'being self-identical'. According to Cowling, the status of these properties is controversial. On an intuitive level, he argues, these properties seem non-qualitative. On the other hand, they don't depend on any specific individual: thus we have good reason to identify them as qualitative.

According to my account, structural properties are qualitative in nature. In particular, the two properties we have taken as examples would be luminously regimented in $\mathcal{L}$ as the formulas '$\exists y(x \neq y)$' and '$x = x$' respectively, and it is easy to see how, for any interpretation $\mathcal{I}$, the truth-conditions of the formulas '$\exists y(c \neq y)$' and '$c = c$' are invariant under $\exists y(x \neq y)$-duplicate worlds and $(x = x)$-duplicate worlds respectively.

5.6.6 Mathematical Properties

Finally, we have *mathematical* properties, like 'being even' and 'having a unique successor', and *species* properties, like 'being H2O'. Cowling argues that the status of both these kinds of properties is controversial. Mathematical properties, he says, are so dissimilar from paradigmatic qualitative properties that it is unclear whether they can be regarded as qualitative. Also, he suggests, it is not straightforward to say that these properties ground qualitative resemblance relations between numbers.

I don't think Cowling's argument is decisive here: despite their dissimilarity with respect to paradigmatic qualitative properties like *mass* and *charge*, mathematical properties might still enjoy the status of qualitative properties. In the end, the entities they apply to are also extremely dissimilar from the entities mass and charge apply to. Furthermore, I find it extremely difficult to work out a clear understanding of what qualitative resemblance between numbers might amount to.

Reflecting these conceptual difficulties, my account doesn't place mathematical properties uniformly across the qualitative/non-qualitative distinction. Take, for example, the property 'having a unique successor'. When evaluated in the context of Peano Arithmetic, a luminous regimentation $\varphi(x)$ of this property is: '$\exists y(Syx) \wedge \forall y \forall z((S(y,x) \wedge S(z,x)) \rightarrow y = z)$'. In this case, the property 'having a unique successor' turns out to be qualitative, since for any interpretation $\mathcal{I}$, the truth-conditions of the $\varphi(x)$ are invariant under $\varphi(x)$-duplicate worlds.

Now consider the property 'being even', and suppose this property is the property 'being divisible by 2 without remainder'. Then, a luminous regimentation of this property in $\mathcal{L}$ is the formula '$M(x, 2, 0)$', where 'M' is a three-place relation that denotes the *modulo* operation. So regimented, 'being even' turns out to be non-qualitative. If instead the property 'being even' is not relational, then a luminous regimentation of it in $\mathcal{L}$ is the formula '$E(x)$'. In this case, 'being even' turns out to be qualitative after all.

It is important to understand that this is not to say that one and the same property, under different representations, can be qualitative or non-qualitative. Remember, in fact, that what counts as a luminous regimentation is a matter of how well the regimentation matches the actual structure of the relevant property. And we can disagree, on the basis of the ontology we endorse, about the actual structure of the property 'being even'.

5.6.7 Species Properties

The same goes for *species* properties, like the property 'being H2O'. Cowling 2015, p. 286 argues that since cases like the Twin Earth suggest that species terms function like proper names, then there is a sense in which we should consider species properties as non-qualitative. The idea is that in the Twin Earth scenario as described in Putnam (1975), the term 'water', and hence the property 'being H2O', seems to be able to distinguish between qualitatively indiscernible entities: Earthly water and Twin-Earthly water. Therefore, Cowling suggests, maybe the property 'being H2O' is indeed non-qualitative.

My intuitions disagree with Cowling's: I don't think Earthly water and Twin-Earthly water are qualitatively indiscernible, even though I agree that they are phenomenologically indiscernible. To see why, imagine to have just one molecule of water (Earthly water, that is) on the palm of your hand. Now, consider the following two properties:

> P: 'Being composed of *one* oxigen atom and *two* hydrogen atoms', and
>
> Q: 'Being composed of *this* oxigen atom and *this* and *that* hydrogen atoms'.

My intuition is that P and Q differ with respect to their qualitative status, and since Q is clearly non-qualitative, P should be considered qualitative. This is why I think that my account correctly characterizes the property 'being H2O' as qualitative, if it is the property 'being composed of *one* oxigen atom and *two* hydrogen atoms'.

As with mathematical properties, I can see ample space for disagreement with respect to the correct structure of species properties. If species properties are non-relational, my framework will still identify them as qualitative. But if the property of 'being H2O' is indeed the property 'being composed of *this* oxigen atom and *this* and *that* hydrogen atoms, or of *this other* oxigen atom and *this other* and *that other* hydrogen atoms, or ... or of *this last* oxigen atom and *this last* and *that last* hydrogen atoms', then my framework would align with Cowling's intuitions and identify 'being H2O' as non-qualitative.

5.7 Conclusion

In this Chapter I have presented a new account of qualitative and non-qualitative properties, and I have argued that it is a promising reductive definition of the qualitative distinction. I have first discussed my account from a philosophical point of view, situating it in the broad non-linguistic tradition of reductive accounts of qualitative properties. Then, I developed a formal framework to distinguish between qualitative and non-qualitative properties. Finally, I argued that my account aligns well with the pre-existing intuitions in the literature about which properties are qualitative and which are not, and provides new valuable intuitions when it comes to the status of those properties about which authors still disagree.

Chapter 6

Reference to Indiscernibles

Now that we have a more precise understanding of indiscernibility and its connection with other notions and theses in Metaphysics and Ontology, we can turn our attention to questions concerning our linguistic practices with respect to indiscernibles. In particular, in this Chapter, we will be concerned with the question: Can we refer to indiscernible entities singularly?

6.1 Introduction

The question of whether we can refer to only one among many indiscernible entities can be found everywhere in the literature about indiscenibles. Numerous authors have attempted to answer it, and virtually every answer so far points toward a negative direction. As for now, almost everyone seems to agree that singular reference to indiscernible entities is impossible. (See, among others, Assadian 2019, Black 1952, and Hellman 2004.)

Here I challenge this view. I argue that the theory of Arbitrary Reference (as developed, among others, in Woods 2014) suggests that there is nothing philosophically problematic in the possibility of singular reference to one among many indiscernibles. Towards my conclusion, I first discuss some relevant literature and propose a starting intuition to the effect that singular reference to indiscernibles is possible (Section 6.1). I then discuss two important philosophical distinctions: the distinction between individuals and non-individuals, and the distinction between metaphysical and epistemic individuation (Section 6.2). One of the ideas behind this Chapter is that the negative answers found in the literature to the question of singular reference to indiscernible entities might depend on two factors: the habit of not distinguishing between indiscernible individuals and indiscernible non-individuals, and some confusion about metaphysical and epistemic individuation, and their role in securing the possibility of singular reference. In Sections 6.3 and 6.4 I discuss various theories of Arbitrary Reference (AR), focusing in particular on the account of AR developed in Woods (2014) and Boccuni & Woods (2020). Finally, in Section 6.5, I show that Woods's account of AR, as well as other similar accounts, is compatible with the possibility of singular reference to indiscernible individuals.

6.1.1 Many Kinds of Indiscernibility

In the previous Chapters we have seen that we can define various relations of world-indiscernibility (henceforth: indiscernibility for short) on the basis of the properties we are interested in. We can say that two entities are *intrinsically indiscernible*, for example, whenever they agree with respect to all their intrinsic properties. (One example of intrinsic indiscernibles are Lewis's (1986, p. 62) duplicates.) Or we can say that two entities are *spatio-temporally indiscernible* whenever they agree with respect to all their spatio-temporal properties — like in certain cases of co-location.[1] More generally, for any set S of properties, we can define a relation of S-indiscernibility:

> ***S*-Indiscernibility (Definition)**: For all entities x and y, x and y are S-indiscernible whenever: for every property P in S, x has P if and only if y has P.

Formally, we can define a collection of indiscernibility relations $\{\equiv_n \colon n \in \mathbb{N}\}$ and provide a definition of indiscernibility for each set S_n of properties and relations via the formula:

$$x \equiv_m y \leftrightarrow \forall P_m(P_m x \leftrightarrow P_m y)$$

where P_m is a property in S_m. Call this formula ID (for Indiscernibility Definition). Virtually all definitions of indiscernibility one finds in the literature are particular instances of ID, with respect to the relevant set of properties. (Or, at least, they are particular instances of a version of ID which also includes relations. It should be easy to see how to generalise ID in this way — however, for simplicity, I will here consider the simpler monadic version of ID. Everything I say about ID can be applied to its polyadic version too.)[2]

It is interesting to note that ID is an open formula. Therefore, where $\mathcal{M} = \langle M, v \rangle$ is a model of our theory of reality and s a suitable variable

[1] For more on co-location, see Smid (2021).

[2] See, among others: Adams (1979), Cowling (2015), French (1989), Hawley (2009), Lowe (2016) Rodriguez-Pereyra (2006), Saunders (2003, 2006), and Wüthrich (2009).

assignment, ID is true in $\mathcal{M}$ whenever $s(x)$ and $s(y)$ are some m_1, m_2 in M, and the relevant bi-conditional holds true for m_1 and m_2. That is: for any $N \subseteq M$ such that $N = v(P_n)$ for some predicate 'P_n', $m_1 \in N$ if and only if $m_2 \in N$. The variable assignment s is a function assigning every individual variable to some entity in the domain, and every predicate variable a subset of the domain. (Recall that, for simplicity, we are only considering unary predicates. Nothing hinges on this.) Furthermore, a formula with a free individual variable 'x' can be assigned a truth-value in some model $\mathcal{M}$ only if s is defined for x according to $\mathcal{M}$. This facts will be useful for the discussion in Sections 6.1.3 and 6.1.4.

6.1.2 Is Reference to Indiscernibles Impossible?

The question of whether we can refer to indiscernible entities has been addressed by numerous authors. In Black (1952, p. 156), for instance, we find the following:

> How can I [consider only one of my spheres and designate it as 'a'], since there is no way of telling them apart? *Which* one do you want me to consider? [...] I don't know how to identify one of two spheres supposed to be alone in space and so symmetrically placed with respect to each other that neither has any quality or character the other does not also have.

Similar concerns echo across the literature. Talking about the Cardinal Four Structure, i.e. a structure consisting only of four distinct places and no structural relations, Hellman (2004, p. 572) asks:

> How is it that any [of the places in the Cardinal Four Structure] is distinct from any other? Indeed, how can we make sense of referring to any one of them as opposed to any other, or mapping any one of them to or from anything else [...]?

Thoughts of this sort are widespread, and one of the reasons why they are so compelling is that it is commonly thought that some kind of individuation is necessary to secure singular reference. Black and Hellman can be read as suggesting that singular reference is possible only when there is some fact of the matter about which entity is *which*, among a certain plurality of entities. And, many would say, indiscernible entities seem to lack this feature.

In this Chapter, I won't argue against that idea that individuation and singular reference are strictly connected. Rather, I will argue against the idea that singular reference is possible only in the case *we* can indeed identify the entities we want to refer to among a plurality of entities. To this end, I will follow Lowe (2003) in distinguishing between two kinds of individuation: *epistemic individuation*, which occurs whenever a subject successfully singles out an individual as a single object of thought, and *metaphysical individuation*, which is a metaphysical relation holding between distinct entities: the *what makes an entity exactly the entity that it is.* (I will expand more on this distinction in the Section 6.2.2.)

In particular, I will think of the connection between individuation and singular reference as follows: singular reference is possible *only* if the entities we want to refer to are individuated in the sense of metaphysical individuation — if there is something which makes these entities what they are and not anything else. Whether we succeed in individuating them (i.e. singling them out) or not, I believe, doesn't bear any consequence on the possibility of singular reference.

6.1.3 An Initial Intuition

We have seen that both Black (1952) and Hellman (2004) hold that singular reference to indiscernible entities is impossible. More precisely, they seem to hold the following three statements: (1) indiscernible entities exist, (2) the relevant definition of indiscernibility is an instance of ID, and (3) indiscernible entities cannot be referred to singularly.

That Black and Hellman hold statement (1) seems uncontroversial. Arguing against the Identity of Indiscernibles, Black (1952) must hold true that distinct indiscernible entities are at least possible. And even if Hellman (2004) is using the Cardinal Four Structure (CFS) as a lever against *ante rem* structuralism, his argument against this strand of structuralism works on the assumption that CFS exists and looks like as described by the *ante rem* structuralist: a structure with only four places, which the lack of structural properties and relations renders trivially indiscernible. Therefore, insofar as their arguments go, both Black (1952) and Hellman (2004) are committed to the existence of indiscernible and yet distinct entities. They also hold (2): Black (1952, p. 155–156) explicitly suggests that his spheres are indiscernible in virtue of sharing all their qualitative properties, and Hellman (2004, p. 570) specifies that the places in CFS are indiscernible in virtue of bearing the same intra-structural relations. And these are just two instances of ID. Finally, it is a cornerstone of both their arguments that these indiscernibles cannot be referred to singularly, which is statement (3). In line with in line with Bach (1987) and Kripke (1980), I take (3) as meaning that there is no reference function which connects our linguistic expressions to these indiscernible entities.

Now, it seems to me that there is a sense in which (1)-(3) are incompatible, if we want our definition of indiscernibility to be non-vacuous: i.e. if we want whatever instance of ID we believe in to be true not only because its left-hand side and right-hand side are false.[3] To see this, take any instance of ID. If we believe that this instance of ID is a correct definition of indiscernibility, we must want it to be true in every model of our theory of reality (whatever theory this is). Furthermore, since we believe that indiscernibles exist, we must believe that the intended model $\mathcal{M}$ of our theory is one which includes indiscernibles and in which the chosen instance of ID holds true.

For this second condition to be the case, however, the relevant instance of ID must be true in $\mathcal{M}$ under every variable assignment. Furthermore, for it to be non-vacuous, the relevant variable assignment from individual variables in our language to the entities in $\mathcal{M}$'s domain must be defined for $\mathcal{M}$'s indiscernibles. Otherwise, 'x' and 'y' being assigned to discernible entities, our definition of indiscernibility would be true only because its left-hand side and its right-hand side are both false: i.e. it would be true, but

[3] Here's an example to clarify the notion of vacuity I have in mind. Suppose we believe that everything is extended. Then, we hold somehow that a good theory of reality must contain the formula '$\forall x Ex$', where 'E' denotes the property 'being extended'. Now, consider the following two cases: (i) the intended model of our theory of reality has an empty domain, and (ii) the intended model of our theory of reality has a non-empty domain, and the domain is the interpretation of the predicate E. Clearly, in both cases our formula holds true. However, only in case (ii) I would say that it holds true in a non-vacuous way.

vacuously so. Therefore, there must be a suitable variable assignment s such that $s(x) = m_1$ and $s(y) = m_2$ for some indiscernible m_1 and m_2 in $\mathcal{M}$. But if this is the case, then we can think about this variable assignment as a reference function, that is: we can say that s specifies m_1 as the referent of 'x' and m_2 as the referent of 'y'. From this, it seems to follow naturally that singular reference to indiscernible entities is indeed possible.

6.1.4 Some Objections

I can see different ways to challenge this intuition on the basis that ID is not quite a good definition of indiscernibility. Here I will discuss two of them. First, one might say that we should take the universally quantified version of ID as our definition of indiscernibility:

$$\forall x, y(x \equiv_m y \leftrightarrow \forall P_m(P_m x \leftrightarrow P_m y))$$

If this is the case then, one might argue, we can stop worrying about variable assignments. This might be true, but this quantified version of ID presents its own issues. This is because we can universally instantiate ID for any pair of individual constants or variables in the language. And again, if we don't want ID to be trivial, we must hold that the intended model of our theory of reality must be some $\mathcal{M}$ with indiscernible elements m_1 and m_2 and an interpretation function υ such that there are some constants 'a' and 'b' in our language such that $\upsilon(a) = m_1$ and $\upsilon(b) = m_2$ (or a suitable variable assignment s and some variables 'x' and 'y' in our language such that $s(x) = m_1$ and $s(y) = m_2$). But then again, we have a function from the constants and variables in our language to the entities in our domain such that indiscernible entities are the unique referents of some linguistic expressions.

Alternatively, one could hold that what I'm mistaking for variables in ID are indeed best interpreted as meta-variables ranging over individual constants. That is: ID should be a schema, rather than a formula. Then, a working version of ID would look like the following:

$$\alpha \equiv_m \beta \leftrightarrow \forall P_m(P_m\alpha \leftrightarrow P_m\beta))$$

Now, one could hold, our definition of indiscernibility can be meaningful even if it doesn't in fact refer to any individual: after all, meta-variables are not supposed to refer to objects in our domain(s). Unfortunately, even this solution won't work. Schemes are in fact commonly interpreted as *sets* of sentences in the object language. Therefore, this new version of ID is just a shortcut for the infinite series of sentences:

$$a \equiv_m b \leftrightarrow \forall P_m(P_m a \leftrightarrow P_m b)$$

$$a \equiv_m c \leftrightarrow \forall P_m(P_m a \leftrightarrow P_m c)$$

$$b \equiv_m c \leftrightarrow \forall P_m(P_m b \leftrightarrow P_m c)$$

$$\ldots$$

However, in any of these sentences, we have individual constants, and not meta-variables. And unlike meta-variables, individual constants are usually thought of as referential linguistic entities. Furthermore, the reference of individual constants is usually defined via the interpretation function, which means that, provided the function is well-defined, each constant for which the interpretation is indeed defined has a unique referent. Therefore, the above schema is meaningful only if all of the sentences it stands for have a truth-value. And any of these sentences, taken at face value, can be true in a non-vacuous way only if its constants refer to some indiscernibles. (Remember that we are reasoning on the assumption that indiscernibles exist.) But if this is the case, then we must admit that it is indeed possible to individually refer to indiscernible entities.

There are many ways to avoid these issues. One way is to adopt some instance of our quantified version of ID whilst rejecting Universal Instantiation, that is:

$$\forall x F(x) \Rightarrow F(a).$$

This is possible, for example, in free logics, which allow for empty domains and non-denoting individual constants. Another way would be to suggest that the quantification in the quantified version of ID should be understood as a form of plural quantification, and that all talks about indiscernibles are indeed inherently 'plural' talks. These are perfectly nice suggestions, and I'm open to accept any of them if some reason can be provided why it is better to interpret our definition of indiscernibility in logics other than classical second order logic. However, there's currently no debate about this issue, and all the authors who talk about indiscernibles seem to do just fine with classical second order logic. (See, for example, Button (2017) and Ladyman et al. (2012).) Furthermore, note that going for logics other than classical first and second order logic *because* no reference to indiscernibles can be obtained would be, at this point, quite circular and unacceptable. For whether we can have such singular reference is just what's at stake here.

This is of course not meant to be a definitive argument, and it shouldn't be considered to have any more force than that of an intuitive observation. In what follows, I will deal with the question whether genuine singular reference to indiscernible entities can be achieved, and if my answer will turn out to be correct, then there will be enough reasons, I hold, to stick with second order classical logic for most of our talks about indiscernibles, even if this will require some revisions on how we think of, and build, models.

In particular, I will try to answer the question of singular reference to indiscernibles throughout the machinery of the so-called Arbitrary Reference (AR), which is, roughly speaking, the idea that we can refer arbitrarily to individual objects, even if we are not able to individuate them uniquely. (See, among others, Breckenridge & Magidor (2012).)

Before discussing Arbitrary Reference, however, I will discuss two distinctions which are crucially important and yet almost never considered within the philosophical literature on indiscernible entities. These are: the distinction between *individual* and *non-individual indiscernibles*, and the distinction between *epistemic* and *metaphysical individuation.* I shall consider them in turn in the next Section.

6.2 Some Important Distinctions

In a recent work, Assadian (2019) argues that the scepticism with which the majority of authors look upon absolutely indiscernible entities (that is: entities that cannot be distinguished by means of non identity-involving properties or relations) is ill-founded. (See Section 1.3 for a discussion of the many reasons of scepticism towards indiscernibles.) In particular, Assadian suggests that there is no metaphysical, epistemological or semantical issue that is unique of this kind of entities. All the concerns that have been raised in the literature about absolutely indiscernible entities, he argues, equally apply to weakly indiscernible entities: i.e. entities which are only discerned by symmetric and irreflexive relations. And since weakly indiscernible entities are philosophically harmless, there is no principled reason to look at absolutely indiscernible entities with any suspicion. In this Section, I take Assadian (2019) as a case study, a representative of the contemporary metaphysical and meta-philosophical literature about indiscernibles.

In line with Black (1952) and Hellman (2004), Assadian (2019, p. 2559) suggests that singular reference to absolutely indiscernible entities is impossible:

> There is [...] nothing which determines the references of the terms standing for [absolutely indiscernible entities], and so [Hellman (2004)] rightly points out that we cannot make sense of referring to any one of them as opposed to any other. [...] It is surely true that we never refer to [absolutely] indiscernible entities by using *singular terms*.

In line with the literature, Assadian (2019) talks about indiscernible entities as if they all belonged to the same ontological kind. In other words, he talks of indiscernibles in the plural, under the implicit assumption that such a unified view of indiscernibles is possible.[4] In what follows, I argue that

[4]Remember that Assadian (2019) is here taken as a case study. Assadian is definitely

such background view is philosophically problematic, for it entails that (1) all the questions that we can meaningfully ask about indiscernible entities of a given kind are equally applicable to indiscernibles of other kind(s), and (2) even when they indeed apply to more than one kinds of indiscernibles, their answer doesn't depend on the kind of the indiscernibles in question. We will see that (1) and (2) are indeed wrong.

6.2.1 Individuals and Non-individuals

The two kinds or categories of indiscernibles I want to distinguish are: indiscernible individuals and indiscernible non-individuals. Indiscernible individuals are individuals which have all their properties of some given kind in common, while indiscernible non-individuals are non-individuals which have all their properties of some given kind in common.

Following some literature in the Philosophy of Quantum Mechanics, I take the main difference between individuals and non-individuals to be that while identity and distinctness apply to individuals, neither identity nor distinctness apply to non-individuals.[5]. That is: suppose x and y are individual entities. Then, the sentences "x is identical to x" and "x is distinct from y" are meaningful. Suppose now that x and y are non-individual entities. In this case, the sentences "x is identical to x" and "x is distinct from y" are meaningless, for identity doesn't apply to non-individuals, and therefore

not alone in making the assumption that indiscernibles belong to a homogeneous category. To my knowledge, distinctions between distinct ontological categories of indiscernibles are present only in some literature in Philosophy of Quantum Mechanics, among which a notable example is French & Krause (2006). The only exception to this trend is represented by Lowe (2016).

[5] See, among others: French & Krause (1995) and French & Krause (2006)

it is a category mistake to say that x is self-identical, or that there is some non-individual which is distinct from x.

This distinction is particularly salient in the Philosophy of Quantum Mechanics, where it is argued that some elementary particles in some physical states are non-individuals in the sense just specified. Suppose, for example, that x and y are two entangled electrons. According to the Received View of Quantum Mechanics, it is a consequence of Quantum Statistics that the relation of identity doesn't apply to x and y, and therefore that the sentences "x is identical to x" and "x is distinct from y" are meaningless. I have given an argument for this point in Section 4.4.

This distinction between individuals and non-individuals looks similar to a distinction suggested by Lowe (2016). Lowe (2016, p. 50) defines an individual as something which obeys two conditions: it has a determinate identity and it counts as *one* entity. Anything that doesn't meet these requirements is then a non-individual. As Lowe remarks, there are three ways in which an entity might fall short of being an individual according to his account: (1) it might fail to have a determinate identity while still counting as one entity, (2) it might still have a determinate identity while not counting as one entity, and (3) it might fail in both having a determinate identity and in counting as one entity.

An example of (1) is, according to Lowe, any of the two orbital electrons of an helium atom. In this case, Lowe (2016, p. 50) suggests, there are definitely *two* electrons, for one is in spin-up state while the other is in spin-down state. However, there is no fact of the matter as to *which* electron is in spin-up state and *which* electron is in spin-down state.

An example of (2) are pluralities: i.e. things which are *many* instead of *one*. *The planets of our Solar System* is a plurality, *the kings of ancient Rome* is a plurality, etc. These entities don't count as one, but as many: the planets of our Solar System are *eight* in number (please don't tell Pluto), while the kings of ancient Rome count as *seven* entities. It is important to note that, for Lowe, *the planets of our Solar System* are not the same as the *set* whose only members are the planets of our Solar System, for this second entity, the set, definitely counts as *one*.

Another quite different example of (2) is stuff. While pluralities fails to count as one entity because they count as more than one, stuff fails to count as one entity because it lacks number altogether. Lowe's (2016, p. 51) example is about *the water in his bathtub*. While it has a determinate identity, for the sentence "the water that was is Lowe's bathtub yesterday is now in the river Thames" is clearly intelligible, the water in his bathtub lacks number, for the question "how many water was there is Lowe's bathtub yesterday" is just meaningless. So there are at least two distinct kinds of entities (i.e. pluralities and stuff), which fall short of being individuals because they fail to count as one entity.

Lowe doesn't suggest examples of (3), and limits himself to the remark that it is indeed difficult to think about entities, either actual or possible, which lack precise identity and do not count as *one*.

The careful reader will have noticed that the distinction I propose between individuals and non-individuals is quite different from the distinction suggested by Lowe (2016), despite a quite unfortunate overlap in terminology. Lowe never considers entities to which identity doesn't apply: even in the case of electrons, Lowe (2016, p. 52) believes that although there is no fact of the matter about which electron is which, among the electrons orbiting around a helium's nucleus, it is still the case that each of those electrons is identical to itself, and that there are two distinct (in the sense of non-identical) electrons orbiting around the helium's nucleus. More generally, Lowe seems to believe that the relation of identity applies across the board.

This intuition is contrary to that of the Received View of Quantum Mechanics, according to which some elementary particles in certain physical states are entities for which "the relation of identity $a = a$ does not make sense". (French & Krause 2006, p. 248) And this last intuition is what I want to capture: that there are some entities which lie beyond the realm of identity.

I want to stress that the philosophical significance of this distinction between individuals and non-individuals, which I used in Chapter 4 to argue that Haecceitism follows from the fact that a version of PII restricted to ordinary spatio-temporal entities is not necessarily true, and which I will use again in this and the following Chapters, is independent from the success or failure of the Received View as the correct interpretation of the metaphysical status of elementary particles. The Received View has given us reasons to think that the relation of identity is not universal. And although these reasons are compelling, they might turn out to be false for quantum particles. However, this doesn't exclude the coherence of the idea that there might be entities which identity cannot reach, as it were. What the Received View has given us, apart from a picture of the actual world according to which identity doesn't apply to (some) elementary particles, is a consistent theoretical framework in which entities without identity exist. It has, in other words, made very difficult to deny at least the possibility of such entities, since it has shown that no contradiction arise from their existence.

6.2.2 Epistemic and Metaphysical Individuation

Another distinction which will be important for what follows is the one between two distinct kinds of individuation. Drawing on Lowe (2003), we can distinguish between an epistemic kind and a metaphysical kind of individuation. Epistemic individuation is the act of 'singling out' an individual entity among a plurality, "[...] as a distinct object of perception, thought, or linguistic reference." (Lowe 2003, p. 75) It is a cognitive effort, which requires an epistemic connection of some sort between the subject of the cognitive endeavour and the entities they will suitably individuate. (It requires, in an old-fashioned sense, the encounter of a mind and a world, or part thereof.) Distinct intelligent beings will sometimes differ in the number and kinds of things they will individuate in some specific situation. Usually, the different the categories and kinds they are familiar with, the different will be the entities they will 'single out' as distinct.

On the contrary, metaphysical individuation is an ontological relation between entities. In this sense, for any individual entity, there is something (some fact of the world, maybe, or some relation or property) which makes that entity exactly the entity that it is, as opposed to any other entity. We say that this something (be it a fact, a relation, or what have you) is what 'individuates', in the metaphysical sense, an individual.

According to Lowe, metaphysical individuation is a necessary condition for epistemic individuation, in the sense that we can single out some entity x as a distinct object of perception or thought only if x is indeed there for us to individuate, that is: if indeed there is something in the world which metaphysically individuates x as a single entity.

6.2.3 Some Important Distinctions Applied

As we have already mentioned, for many authors the issues of individuation and individual reference go hand in hand. One clear example is the passage from Black (1952) quoted in Section 6.1.2. The dialectic of this passage, starting few lines before, is the following. After laying out in full details the description of a symmetric world containing only two indiscernible spheres, Black is asked by his imaginary antagonist to consider only one of the two spheres in his world and call it 'a'. (Famously, Black 1952 is written in the form of a dialogue between Max Black, 'B', and an imaginary interlocutor, 'A'.) Black answers that since there is no way of telling the spheres apart, the request of considering just one of them and giving it a name is indeed impossible. His antagonist rejoins by suggesting that, upon being asked to "pick any book off the shelf", it would be indeed foolish of him to answer "which one?".

Despite agreeing, Black points out that there's a fundamental difference between the two scenarios: while in the case of the books he *knows* how to identify a book among many in a shelf, that is not so in the case of the spheres. We can therefore understand Black as thinking at the difference between the books and the spheres as a difference in the possibility of epistemic individuation. What Black is doing is to relate the possibility of successful individual reference to the success of some form of epistemic individuation: since he knows how (and therefore can) individuate any book among a plurality of books in a shelf, there is no problem in calling some book 'a' and some other book 'b'; however, since he doesn't know how (and hence cannot) individuate any of the two spheres in his symmetrical universe, then the act of naming one of them 'a' and the other 'b' is necessarily unsuccessful.

That Black (1952, p. 156) is only concerned with epistemic individuation is clear from the fact that he considers his thought-experiment as a counterexample to Strawson's version of PII, according to which no two entities can have all their qualitative properties in common. That this fact alone entails the necessity of some metaphysical individuation of the two spheres can be best understood if we ask what it means for some plurality of entities to run against PII. (For a similar reasoning, see Section 4.5.) PII is in fact usually regimented as the following second order sentence:

$$\forall x, y(\forall P(Px \leftrightarrow Py) \rightarrow x = y)$$

where 'x' and 'y' are individual variables and 'P' is a second order variable standing for qualitative properties. When we consider this regimented form of the principle, it is easy to see that objects a and b disobey PII if and only if:

· $\forall P(Pa \leftrightarrow Pb)$, and

· $(a \neq b)$.

What this means is that Black must believe that indeed his spheres agree with respect to all their qualitative properties, and that they are numerically distinct in the sense of being such that $(x \neq y)$. But this last fact has many interesting consequences.

First, that the spheres Black is talking about are individuals in the sense specified in Section 6.2.1. (For recall: if they weren't individuals, to say that they are distinct, in the sense of non-identical, would simply be meaningless.) Second, that each sphere must be identical with itself, for if x is not an inconsistent entity and identity applies to x, then by the Reflexivity of Identity it follows that x must be identical with x.

But since any of Black's spheres is indeed identical with itself and distinct with respect to all the other things in the scenario, then they are already metaphysically individuated, even if only by the weakest form of metaphysical individuation. The question of what exactly individuates them is not particularly relevant here, and it would depend on how structured is one's conception of individuality. Notice though that under a very minimal conception of individuality, we could say that the mere fact that any sphere is identical to itself and distinct from any other sphere is what individuates it in the relevant sense. Sure, this might be taken to entail that identity is fundamental, but this should be expected, if we indeed believe, with Black, that there are entities which are qualitatively indiscernible.

Let's now consider Hellman's (2004, p. 527) concern, as reported in Section 6.1.2. It is important to remember that Hellman (2004) is talking about mathematical entities from a structuralist point of view, according to which mathematical entities are individuated by the structural relations they stand in to all the other entities in a certain structure.

As an example, consider the natural number 1 and the Natural Numbers Structure. According to the structuralist, what individuates the number 1 in a suitable structure (if we are considering the natural numbers structure, any ω-sequence would do) are the relations it stands in to the other elements of the structure: the numbers 0, 2, 3, 4, 5, etc. In particular, the fact that 1 is the successor of 0 and the predecessor of 2 (or, if you want, the fact that 1 is the only successor of some number that is no number's successor) is part of the constitution of 1, and together with the other relations of the same kind, it individuates 1.

Hellman (2004) takes issue with those mathematical entities which are indiscernible within a given structure. And to deliver his point, he discusses the Cardinal Four Structure: a structure composed only of four places with no structural relation. Now, Hellman argues, it is impossible to refer to only one of the places in the Cardinal Four Structure, as opposed to any other. For "how is it that any [of these places] is distinct from any others?" (Hellman 2004, p. 572)

These structures, Hellman remarks, are something the *ante rem* structuralists should worry about, if it's true that, as Keränen (2001) and Burgess (1999) argue, *ante rem* structuralism entails the Identity of Indiscernibles.[6]

I won't take issue with *ante rem* structuralism nor with any other struc-

[6]The argument, formulated independently by Keränen (2001) and Burgess (1999) can be summarized as follows: there is a sense in which, according to the *ante rem* structuralist, mathematical entities are sufficiently identified by the relations they bear with the other entities in the relevant structure. But if this is true, then the identity of those entities depends on such relations, and hence no two entities can be in exactly the same relations with respect to the exact same entities in a given structure. And this is just a version of PII restricted to structural properties.

turalist position in the Philosophy of Mathematics, and the following argument should not be taken as a defence of some form of structuralism against some other form of it. For the present purposes, I'm only interested in the fact that according to some well understood theories of (mathematical) entities, there are (abstract) entities which are indeed indiscernible, and about which the question about whether we can identify them and refer to them have been asked.

It is important to note that, at least at a first sight, Hellman's (2004) concern seems different from Black's (1952). Hellman seems less interested in how we can individuate the members of the Cardinal Four Structure, and more interested in the question whether there is some sense in which we can say that there's indeed something that individuates them. In other terms, Hellman's concern seems to be with metaphysical individuation rather than epistemic individuation, and how this affects the possibility of reference.

I have no idea whether there is some fact of the matter which individuates the members of the Cardinal Four Structure. However, we can run a case by case argument, and see what we can conclude from examining all the possibilities. We begin by noticing that Hellman seems convinced that some version of PII does indeed follow from *ante rem* structuralism.

Does this alone tell us anything about the category to which mathematical entities belong? The answer is, of course, in the negative. However, if we assume, as Hellman seems to do, that the relevant version of PII is not vacuously true, then we can conclude that at least some of the entities the structuralist is talking about are indeed individuals. If PII is not vacuously true, in fact, then must be entities which non-vacuously satisfy it. And where there are entities which non-vacuously satisfy PII, there are individuals.

This of course doesn't tell us anything about the places in the Cardinal Four Structure yet. However, Hellman suggests that the places in CFS violate PII: and if he is right, then the same argument we rehearsed against Black's intuition can be used once again. If these entities violate PII they must be individuals. But if they are individuals, then they must be individuated. Hellman argues that only structural relations can individuate mathematical entities, and in the case of the Cardinal Four Structure there is no structural relation in play. However, as Hellman acknowledges, one can say that despite there being no structural relations in play, the relations of identity and numerical difference still holds for the places in the structure considered (this is suggested to him by Shapiro in correspondence). And if identity is in play, we can once again resort to some minimal account of individuation where the relations of identity and numerical difference indeed metaphysically individuate the objects at issue.

Once again, this means resorting to a primitive notion of identity, but this should be expected, in cases in which PII is violated. If this kind of minimal individuation is possible, we will see, the doctrine of Arbitrary Reference will give us enough theoretical framework to be able to suggest that we can indeed singularly or individually refer to the entities in the Cardinal Four Structure.

Let's take stock. When it comes to indiscernible entities, the issues of the possibility of their individuation and the possibility to successfully refer to them by means of singular expressions are often discussed together. This suggests that these issues are somehow intimately related. Lowe (2003) distinguishes two kinds of individuation: epistemic and metaphysical individuation. He defines the first as the cognitive act of singling out an entity from a plurality as a single object of thought, and the second as an "ontological relationship between entities: what 'individuates' an object, in this sense, is whatever it is that makes it the single object that it is". (Lowe 2003, p. 75)

We have seen that despite metaphysical individuation being necessary for epistemic individuation (in the sense that we cannot single out what's not there to be singled out), it is not sufficient. In the case of Black's spheres, we have metaphysical individuation without epistemic individuation — and this is the case even in the Cardinal Four Structure, on the assumption that the entities in it are individuals and we stick to a minimal notion of individuation.

I agree with Black (1952) and Hellmann (2004) that the possibility of singular reference goes hand in hand with the possibility of individuation. Unlike Black (1952), however, I suggest that the relevant notion of individuation at issue is the metaphysical notion, not the epistemic one. In particular, I hold that individual reference is possible whenever the objects of reference are suitably metaphysically individuated. The fact that I, as an intelligent being, am not capable of singling out any of them as a single object of perception or thought doesn't impact on my ability to refer to only one of them, as opposed to any other, with a singular term.

In what follows I argue that the theory of Arbitrary Reference provides a suitable formal framework in which to understand how singular reference to indiscernible individuals, like Black's spheres and the places in the Cardinal Four Structure, can be obtained.

6.3 Arbitrary Reference

The key idea behind Arbitrary Reference is that, in some contexts, we can refer to individual entities with some degrees of arbitrariness. One of the examples that are often used to introduce the thesis is that of a mathematician beginning a proof with the words: "Let n be an arbitrary natural number". Later in the proof, our mathematician will write sentences like, say: "[...] and if n is indeed greater than 2, then *it* must be greater than the smallest prime [...]". The relevant questions for the doctrine of Arbitrary Reference are: What is the mathematician talking about when he talks about n? And what is the linguistic nature of this term 'n'?

It seems natural here to say that 'n' is a proper name referring to some natural number. Two reasons for this. First, the surface grammar of the mathematician's sentence is akin to the grammar of sentences like, say, "Let John be the average Cambridge student", and "Let Dedekind be the person who first formulated the Peano Axioms" — and all these sentences seems to be referential expressions involving proper names. Secondly, our mathematician seems to use 'n' exactly as a proper name: he seems to use it to 'name' some number, to which he then refers back, even anaphorically, throughout his entire proof. Furthermore, it seems that the purpose of sentences like "[...] and if n is indeed greater than 2, then *it* must be greater than the smallest prime [...]" is indeed that to state true facts about n. By treating 'n' (and other terms of this kind, called 'instantial terms') like any other ordinary name, Arbitrary Reference makes good of this initial intuition.

There are several accounts of Arbitrary Reference, and they all fall into one of two distinct types, according to how they characterise the kind of reference in question. Accounts of type 1 characterise this kind of reference as non arbitrary reference to arbitrary individuals. Accounts of type 2 hold that the reference at issue is best understood as arbitrary reference to non arbitrary individuals. So the main difference between type 1 and type 2 accounts is about where the arbitrariness of expressions like the one used by our mathematicians should be located.

6.3.1 Arbitrary Individuals

The most systematic account of type 1 is the one proposed by Fine (1983) and later developed in Fine (1985a) and Fine (1985c). For Fine, instantial terms work just like the more common proper names, and refer to 'arbitrary individuals'. Arbitrary individuals are abstract entities, like sets or propositions, and they exist only in an 'ontologically neutral' sense. Fine (1983, p. 56) explains this with an example. There is a sense in which a nominalist about numbers holds that numbers do not exist. However, the same nominalist can agree that there are numbers in another sense, by agreeing that, say, the sums of two primes is not a prime.

Arbitrary individuals are associated with suitable ranges of non arbitrary individuals, which act as their values. For example, the range associated to the 'arbitrary natural number n' is the set $\mathbb{N}$ of natural numbers, and the range associated to the 'arbitrary English man' is the set of all English men. Fine's original theory regiments the attribution of properties to arbitrary individuals by means of the following principle:

For any *generic* condition φx, φa is true iff $\forall i \varphi i$ is true.

Here, 'φ' is any generic condition, 'a' is the name of some arbitrary individual, and 'i' is a variable ranging over all and only the non arbitrary individuals in the range of a. Fine (1983) argues that this principle only applies to generic conditions, and cannot be applied to 'classical' conditions.

According to Fine, generic conditions include ordinary predicates, like for example 'being tall', as well as "[...] all of the conditions obtainable from them by means of of the classical operations of quantification and truth-functional composition." (Fine & Tennant 1983, p. 63) Classical conditions, on the contrary, include predicates like 'being an individual human being' and 'being in the range of', as well as all the conditions that can be obtained from them.[7]

Fine suggests that the distinction between generic and classical conditions has consequences for the semantic role of the instantial terms in sentences like "Let n be a natural number". In particular, if our reading of the predicate 'being a natural number' is classical, then the name 'n' is referential in nature, in the sense that it indeed refers to an arbitrary object. On the other hand, in case our reading of the predicate is generic, 'n' serves a mere representational role, in the sense of representing all the individual numbers in the range of some arbitrary number a.

Fine's theory of arbitrary individuals is further complicated by the introduction of other distinctions, for example the distinction between inde-

[7]Fine (1983, p. 65) also distinguishes between a generic reading and a classical reading of certain predicates, like for example 'being a number'. He says: "On a generic reading, [this predicate] is inclusive of all arbitrary numbers; on a classical reading, it is exclusive of them". Therefore, an arbitrary number a is a number according to the generic reading, but not according to the classical reading.

pendent and dependent arbitrary individuals, and the distinction between 'vacant' and 'occupied' arbitrary individuals. (See: Fine 1985c, p. 75–80.) I won't discuss these further features of the theory of arbitrary individuals here, for I think what I have said so far is enough to understand the underlying idea of what Arbitrary Reference looks like to a friend of arbitrary individuals. Let's then go on and discuss some of the type 2 accounts of Arbitrary Reference which have been proposed in recent years.

6.3.2 Type 2 Arbitrary Reference

Recall that, according to accounts of type 2, Arbitrary Reference is seen as reference to non arbitrary individuals. According to all these accounts, the arbitrariness is located in the mechanism of reference fixing, and not, as Fine suggests, in the nature of the objects referred to. According to all these accounts, when our mathematician says: "Let n be an arbitrary number", he is actually referring, albeit arbitrarily, to one particular number, say: 5, or 10, or 5762.

Breckenridge and Magidor (2012, p. 377) define Arbitrary Reference (AR) as the following thesis:

> **Arbitrary Reference (AR):** It is possible to fix the reference of an expression arbitrarily. When we do so, the expression receives its ordinary kind of semantic value, though we do not and cannot know which value in particular it receives.

This thesis has two components. The first is a meta-semantic component, in the sense that it tells us what the semantic behavior of a given expression is (in this case, an expression containing some instantial term). The second component is instead epistemic: it tells us that it is impossible to come to know which entity we have referred to by a suitable expression, when we have managed to refer to it arbitrarily. The three accounts I discuss diverge with respect to how they answer the following questions:

Q1: What determines the reference of an instantial term in a suitable expression involving AR?

Q2: Why can't we know what is the reference of an instantial term?

6.3.3 Breckenridge & Magidor

According to the account proposed by Breckenridge and Magidor (2012, p. 379), there is nothing which determines the reference of an instantial term in cases of arbitrary reference. The fact that our mathematician has referred to the number 55, say, instead of 29 is ungrounded. There is no non-semantic fact which determines which number the mathematician referred to, apart from the fact that the mathematician has indeed referred to the number 55.

This account of AR flies in the face of all the standard theories of reference, according to which semantic facts are determined by non-semantic facts.[8] Breckenridge and Magidor (2012, p. 379–380) accept this without turning a hair. "We accept" they say, "that AR conflicts with the commonly held view that semantic facts supervene on use facts. [...] [W]e insist that the view that semantic facts supervene on use fact is simply incorrect".

[8]At least, this is so at a first sight. In Chapter 7, I argue that there are ways to counter the challenge, often posed to theories of AR, from free-floating semantic facts.

The fundamentality of some semantic facts is also, according to the authors, the reason why it is impossible to know the referent of an instantial term (in a suitable expression of the kind in question). There is indeed no fact of the world which determines in advance the referent of an instantial term within an arbitrarily referring expression, and there is no non-semantic fact which determine the referent of an instantial term after the utterance of an arbitrarily referring expression has been made.

6.3.4 Enrico Martino

The account of AR proposed in Martino (2001) links the notion of 'arbitrary reference' with that of 'choice act', thought as any act of selecting one alternative over another. Martino develops his account of AR as an extension of Hintikka's (1996) game theoretical semantic for first order logic. The (ideal) agents in Hintikka's semantics, as well as in Martino's account, are endowed with the possibility of choosing every individual in a specific domain of discourse, and of giving it a name. (This is not an axiom of Hintikka's and Martino's theories: rather, it looks more like a supposition, something taken for granted at the outset, and arguably a necessary feature of any *ideal* agent.)

With these notions in the background, Martino (2001, p. 69) explains his take on AR in the following terms. He supposes that we have direct access to some ideal agent, which in turn has direct access to all the individuals in our domain of discourse. When we start our mathematical proof with the supposition "Let n be some arbitrary natural number", we entrust our ideal agent with both the choice of some natural number, and the naming of it by the term 'n'. So it is the ideal agent, and not us, that chooses some number, say 55, and 'baptizes' it with the name 'n'. (This is Martino's answer to Q1 above.)

Once the reference of 'n' is fixed in this way, Martino (2001, p. 69) secures our ignorance of the referent of 'n' by stipulating that the ideal agent will not communicate their choice to us. (Answer to Q2.) In order for his game theoretical theory of AR to explain the role of instantial terms in mathematical proofs with suppositions, Martino (2001, p. 69) has to endorse a principle according to which "[e]very [individual in] the domain of discourse is capable of being chosen by the ideal agent". He calls this principle the 'Choice Act Principle'. Unlike Breckenridge and Magidor, Martino is not committed to the thesis that semantic facts don't supervene on non-semantic ones.

6.3.5 Jack Woods

Unlike both Breckenridge and Magidor (2012) and Martino (2001), Woods suggests a *supervaluationist* account of Arbitrary Reference. Developed in Woods (2014) and Boccuni & Woods (2020), this view of AR explains the arbitrariness of instantial terms by associating them to classes of choice functions on the relevant domain of discourse.[9]

The main difference between Woods's account and the others presented in this section is that while both Breckenridge and Magidor (2012) and Martino

[9]I call this the Woods's account because in Boccuni & Woods (2020) the authors clearly explain that Boccuni's approach to AR is far from supervaluationist. There, they discuss two distinct approaches to Arbitrary Reference, championed by Boccuni and Woods respectively. In this and the following paragraphs, I'll be only interested in Woods's one.

(2001) model the referential nature of instantial terms by assigning them individuals in the domain (selected through suitable choice functions), Woods's account assigns instantial terms a class of functions from the power set of the domain to the domain. In Woods's view, in fact, "[...] an arbitrary expression refers *over* the class of objects which would satisfy it if it functioned like a device of canonical reference". (Boccuni & Woods 2020, p. 309)

6.4 Woods's AR in Details

To better understand Woods's account, we have to introduce some formalism. According to Woods, the logical form of expressions containing instantial terms (or more generally: devices of arbitrary reference) is that of formulas containing some logical indefinites. Woods (2014) connects the relevant expressions to logical formulas containing Hilbert's famous ε operator. This is a variable-binding term operator, and can be applied to formulas with free variables to obtain a term.[10]

Consider, for example, the formula '$A(x)$', with free variable 'x'. If we attach Hilbert's operator to '$A(x)$', we obtain the term '$\varepsilon.xA(x)$'. The intended interpretation of '$\varepsilon.xA(x)$' is: 'something, if anything, that satisfies '$A(x)$'; if nothing does, something else'. In case there are entities in our domain that satisfy '$A(x)$', the term '$\varepsilon.xA(x)$' will denote one of them, arbitrarily chosen. In case nothing in our domain satisfies '$A(x)$', the term '$\varepsilon.xA(x)$' will instead denote an arbitrary individual in the domain.

[10]For more on variable-binding term operators, see Da Costa (1980).

6.4.1 Types

According to the framework constructed in Woods (2014), variable-binding term operators denote functions of type $((e \Rightarrow t) \Rightarrow e)$, namely: total functions $f : \wp(D) \rightarrow D$ from the power set of the domain to the domain. To explain this, we have to define type symbols. We do it as follows:[11]

- 'e' and 't' are type symbols; and
- if 'S_1', ... 'S_n' and 'S' are type symbols, then so is '$(S_1, ...S_n \Rightarrow S)$'.

Given a domain of quantification D, we interpret the type symbols as follows:

- we let 'e' denote D;
- similarly, we let 't' denote the set of truth-values $\{T, F\}$;
- finally, we let '$(S_1, ...S_n \Rightarrow S)$' denote the set of functions from the cartesian product $S_1 \times ... \times S_n$ to S.

Equipped with the vocabulary of type symbols, we can better understand what is the type the functions associated with 'ε' belong to. We have said that a variable-binding term operator like 'ε' denotes functions of type $((e \Rightarrow t) \Rightarrow e)$.

Given the intepretation specified above, we know that the type-symbol '$(e \Rightarrow t)$' denotes all the functions from our domain D to the set of truth-values $\{T, F\}$. Since each of these function will associate either T or F to the elements of the domain, we can interpret each of these function as a

[11] This definition, and the following interpretation, are taken from Woods (2014, p. 280).

determinate subset G of the domain D. In particular, given a function f of type $(e \Rightarrow t)$, we define G as the set $\{d \in D : f(d) = T\}$.

Since we are considering the set of all such functions, we can interpret $(e \Rightarrow t)$ as the power set of our domain. It follows that functions of type $((e \Rightarrow t) \Rightarrow e)$ are functions from the power set of D to D, such that to each subset G of D, these functions associate a particular element d in D.

6.4.2 Choice Functions

Given a model $\mathcal{M}$ with domain M, one natural way to interpret terms like '$\varepsilon.xA(x)$', where 'A' is some predicate symbol, is the following:

$$\varepsilon^{\mathcal{M}}.xA^{\mathcal{M}}(x) = \begin{cases} \text{some arbitrary } m \in A^{\mathcal{M}}, & \text{if } A^{\mathcal{M}} \text{ is nonempty} \\ \text{some arbitrary } m \in M, & \text{otherwise} \end{cases}$$

where $\varepsilon^{\mathcal{M}}$ and $A^{\mathcal{M}}$ is the interpretation of 'ε' and 'A' in $\mathcal{M}$, respectively. In this framework, the functions associated with expressions like '$\varepsilon.xA(x)$' can be modeled as functions $f : \wp(D) \to D$ from the power set of the domain to the domain.

These functions associate to each nonempty subset of the domain an element of that subset, and to the empty subset one arbitrary element of the domain. That is, they are functions f of type $(e \Rightarrow t) \Rightarrow e)$ such that, for all g of type $(e \Rightarrow t)$:

- $g(f(g)) = T$ if the range of g is not $\{F\}$, and

· $g(f(g)) = F$ otherwise.

It is easy to see that, if we defined these functions as satisfying only the first condition, they would be just classical choice functions as per Zermelo-Fraenkel's Set Theory with Choice (ZFC). However, we need also the second condition to model the semantic behaviour of the ε operator.

With it, the functions we are defining behave like classical choice functions in all the relevant respects, with the sole difference being that unlike choice functions in ZFC, they are also defined for the empty set. This addition is however conservative with respect to the behavior of classical choice functions, and I follow Woods (2014, p. 286) in calling the functions associated with expressions like '$\varepsilon.xA(x)$' (total) choice functions.

6.4.3 Semantics

Now we can delve into the semantics for instantial terms developed in Woods (2014) and Boccuni & Woods (2020). Consider again the supposition "Let n be some natural number", made by our mathematician in their proof. The logical structure of this expression, according to Woods, is a formula containing some variable-binding term operator, like 'ε'. Although neither Woods (2014) nor Boccuni & Woods (2020) offer a suitable logical translation of the above supposition, one natural option according to Woods's approach is to regiment our mathematician's sentence with the formula '$n = \varepsilon.xx \in \mathbb{N}$'. This expression is then associated to the set of all choice functions from $\wp(\mathbb{N})$ to $\mathbb{N}$. All these functions will associate to 'n' some number in $\mathbb{N}$.

According to Woods, if we want our expression to refer arbitrarily, we cannot entrust a single choice function to provide the unique reference of the expression. (As we do when modeling AR in line with Breckenridge & Magidor, and Martino.) The arbitrary character of our mathematician's sentence is preserved only when we consider the set of all choice functions available. The expression will thus refer *over* the set of objects it specifies (in this case: the natural numbers), and unique referents for it will pop out only when we start considering distinct precisifications for the expression at hand.

In particular, each precisification will correspond to a specific choice function, and will thus deliver a unique referent for the instantial term 'n'. Thus, 'n' will refer to the number 1 according to *precisification 1*, say; it will refer to 2 according to *precisification 2*; and so on. When no precisification is selected, the instantial term refers over the class of natural numbers. This doesn't mean, however, that 'n' is not a device of singular reference, explains Woods: for all the functions associated to it are of type $((e \Rightarrow t) \Rightarrow e)$, and thus are functions that provide unique referents for 'n'.

Finally, in Woods' account, a sentence involving an instantial term will be true only when it is determinately true (i.e. true according to every precisification), and false only when it is determinately false (i.e. false according to every precisification). In any other case, it will be neither true nor false. As an example, consider the sentence "[...] n is even", wrote down after the supposition that n be a natural number. Since according to some precisification the number referred to by 'n' is indeed even (think about the precisification where the function's output is the number 2), the sentence is not false. And since there is some precisification according to which the number referred to by 'n' is instead odd (for instance, the precisification where the function's output is 1), the sentence is not true either. In contrast, the sentence "[...] n is greater than or equal to 0" is true, since it is true according to all possible precisifications.

To summarize. According to Woods, the arbitrariness of certain linguistic expressions containing instantial terms is best modeled by associating them classes of choice functions over suitable domains of quantification. These functions provide the referents of the expression under every precisification of it. The existence of these function is what grounds the possibility of arbitrary reference, and thus is the way in which Woods would answer Q1 above.[12] As for Q2, Woods's position is that there is no determinate fact about which precisification is the right one, given a context of utterance. Since there is no such fact, we don't and cannot know which precisification is indeed the one associated with our expression: where there's nothing to know, you can't blame someone for not knowing.

Finally, it is interesting to note that by relaxing Tarski's ([1936] 1983) famous criterion for logicality, Woods's framework allows us to consider logical indefinites (like Hilbert's ε operator and Russell's η operator) as logical constants. (And even though this is not related to the scope of this Chapter, isn't this by itself a pretty cool result?)

[12] Recall that Q1 is the question: What determines the reference of an instantial term in a suitable expression involving AR? On the other hand, Q2 asks: Why can't we know what is the reference of an instantial term?

6.5 Arbitrary Reference Applied

In Section 6.3, I said that accounts of AR fall into two distinct types. Accounts of type 1 see AR as non arbitrary reference to arbitrary individuals, while accounts of type 2 see AR as arbitrary reference to non arbitrary individuals. That is: although both accounts of type 1 and accounts of type 2 agree that expressions like "Let n be a natural number" involve some arbitrariness, they disagree on where exactly the arbitrariness is.

Accounts of type 1 locate the arbitrariness in the kind of individuals referred to be the instantial term 'n': according to them, 'n' refers to a particular kind of abstract entities, namely: arbitrary individuals. Different accounts of type 1 will disagree about the characteristics of arbitrary individuals. However, all of them will agree on the fact that the mechanism of reference-fixing for instantial terms like 'n' in the contexts of expressions like "Let n be a natural number" doesn't involve any arbitrariness. The referent of 'n' is fixed in exactly the same way in which the referent of the name 'John Lennon' in the expression "John Lennon loved Yoko Ono" is fixed. The only difference is that unlike 'John Lennon', which refers to a non arbitrary individual, 'n' refers to an arbitrary individual.

On the other hand, accounts of type 2 don't need to introduce new kinds of entities: they are fine with the ordinary, non arbitrary individuals and the way they are usually characterized. According to these accounts, the term 'n' in the contexts of expressions like "Let n be a natural number" refers to one of the ordinary natural numbers. However, unlike the name 'John Lennon' in the expression "John Lennon loved Yoko Ono", whose referent is fixed via a function which involves no degree of arbitrariness, the way the term 'n' in the contexts of expressions like "Let n be a natural number" is assigned to one specific natural number instead of another involves some degree of arbitrariness. Continuing with our talk of functions, we can say that unlike with the name 'John Lennon' in the expression "John Lennon loved Yoko Ono", the referent of the term 'n' in the expression "Let n be a natural number" is fixed by some choice function.

In what follows, I will not be dealing with accounts of type 1. The question I want to answer in this Chapter, in fact, is not how we are fixing the reference of instantial terms in the contexts of suppositions like "Let n be a natural number", or "Let Pierre be an arbitrary French man". Rather, I want to answer the question whether there is some framework which allows us to claim that in certain contexts we can refer to indiscernible individuals, like the places in the Cardinal Four Structure, or the individual spheres in Black's scenario. Both the places and the spheres, however, are not arbitrary entities.[13] Therefore, it should be easy to see that any account of type 1 would be unsuited for my purposes. What I need is instead an account of type 2, because what I want is a way to claim that with the expression, say, "Let a be one of the spheres", in the context of a discussion about Black's scenario, we can indeed refer to only one of the ordinary, non arbitrary spheres in Black's universe.

In this Section, I show how we can refer to indiscernible individuals via

[13]The question whether there might be indiscernible arbitrary individuals is interesting on its own, but it's not one I will be concerned with here. I suppose the answer to such question would greatly depend on which theory of arbitrary reference we adopt, and on the question of which properties count as qualitative when arbitrary entities are involved.

arbitrary reference by applying the theories we have been discussing in Sections 6.3 and 6.4 to the cases of Black's (1952) indiscernible spheres and Hellman's Hellman (2004) places in the Cardinal Four Structure. In doing this I will mostly focus on Woods's (2014) account of AR, for it is the most complex among the accounts of type 2 we have discussed so far. However, this doesn't mean that we can refer to indiscernible individuals only through Woods's framework. It is an easy exercise to prove that reference to indiscernibles can be obtained within Breckenridge & Magidor's and Martino's accounts too.

6.5.1 The Cardinal Four Structure

Let's then begin with applying Woods's supervaluationist AR to Hellman's CFS. The Cardinal Four Structure is a degenerate case of an *ante rem* structure, a structure where we have only four distinct places, and no structural relation between them.[14] We can represent CFS pictorially as follows:

Figure 6.1: Cardinal Four Structure.

where each dot represents a place, and the absence of arrows represents the fact that there are no structural relations between the places in this structure.

As we saw in Section 6.2.3, the places in CFS are indiscernible individuals, namely: although they share the same structural properties and relations, each place is identical to itself and distinct (in the sense of non-identical)

[14] I have discussed similar structures in Section 1.1.3.

from any other place. Therefore, the places in CFS are metaphysically individuated, even if only by the weakest form of metaphysical individuation.

So can we refer to only one of them, as opposed to any other? The answer seems to be positive. We could for example make the supposition "Let a be one of the places in CFS". According to Woods (2014), this expression has the form '$\varepsilon.xA(x)$', where 'A' denotes the property 'being a place in the Cardinal Four Structure". Therefore, supposing for simplicity that the domain of discourse D is our intended model $\mathcal{M}$ *is* the set of places in CFS (that is: $D = A^{\mathcal{M}}$), our expression gets interpreted as:

$$\varepsilon^{\mathcal{M}}.xA^{\mathcal{M}}(x) = \begin{cases} \text{some arbitrary } d \in A^{\mathcal{M}}, & \text{if } A \text{ is nonempty} \\ \text{some arbitrary } d \in A^{\mathcal{M}}, & \text{otherwise} \end{cases}$$

Following Woods, to this expression we associate the set of total choice functions from $\wp(A)$ to A. Finally, since A is nonempty, we know that any choice function associated to the supposition "Let a be one of the places in CFS" is such that, when its input is A, its output is either one place, or another, or still another, or still another. Therefore, we have a total of four non-equivalent precisifications.

Our supposition will then refer *over* the set of places of the Cardinal Four Structure, and the instantial term 'a' will have unique referents according to distinct precisifications. The fact that there is no determinate fact of the matter as to which is the correct precisification in a given context of utterance doesn't mean, once again, that 'a' is not a device of singular reference. This is because all the functions associated to our supposition are of type $((e \Rightarrow t) \Rightarrow e)$, and thus are functions that provide unique referents for 'a'.

Breckenridge and Magidor's account of AR is much simpler. On the assumption that our domain of discourse is the set of places in CFS, our supposition "Let a be one of the places in CFS" is associated with only one choice function from $\wp(D)$ to D. Whatever the output is of that choice function when the input is D, that will be the unique referent of the instantial term 'a'.

6.5.2 Black's Spheres

The situation with Black's spheres is similar. Much like the places in CFS, we cannot distinguish the two spheres in Black's scenario by means of any qualitative property or relation. However, much like the places in CFS, each of Black's spheres is identical to itself and distinct (in the sense of non-identical) from the other sphere.

So let's make the supposition "Let a be one of the spheres in Black's scenario". Again, Woods's (2014) account dictates that this expression has the form '$\varepsilon.xS(x)$', where 'S' denotes the property 'being a sphere in Black's scenario'. Therefore, supposing for simplicity that the domain of discourse D in out intended model $\mathcal{M}$ *is* the set of spheres in Black's scenario (that is: $D = S^{\mathcal{M}}$), our expression gets interpreted as:

$$\varepsilon^{\mathcal{M}}.xS^{\mathcal{M}}(x) = \begin{cases} \text{some arbitrary } d \in S^{\mathcal{M}}, & \text{if } S \text{ is nonempty} \\ \text{some arbitrary } d \in S^{\mathcal{M}}, & \text{otherwise} \end{cases}$$

Again, our supposition gets associated the set of total choice functions from $\wp(S)$ to S. In the case of Black's scenario there are only four such functions. Two of them output one sphere when input S, the other two output the other sphere when input S. Therefore, since our domain is nonempty, we have a total of two non-equivalent precisifications.

Again, the instantial term 'a' in our supposition "Let a be one of the spheres in Black's scenario" will refer *over* the set of spheres in Black's scenario, and it will have unique referents according to distinct precisifications. That there is no determinate fact of the matter as to which precisification is assigned to which context of utterance doesn't mean, once again, that 'a' is not a device of singular reference. All the functions associated to to our supposition provide in fact unique referents for the instantial term 'a'.

6.6 Conclusion

In this Chapter I suggested that the common intuition according to which it is impossible to singularly refer to only one among many indiscernible entities is mistaken, at least when the indiscernibles at hand are individuals. The main idea of this Chapter was that if identity applies to indiscernibles, then they can be the output of some (total) choice functions, which we can use as suitable reference functions when talking about indiscernibles.

The idea that reference can be sometimes modeled via choice function is at the heart of the theory of Arbitrary Reference, when this is understood as reference to non arbitrary individuals. I have discussed various interpretations of this idea and applied them to the case of indiscernibles. It turns out that if the extant theories of Arbitrary Reference are correct, then there is no philosophical issues arising from the possibility of singular reference to indiscernibles. For as we have seen, Arbitrary Reference can make sense of the idea that the instantial term '*a*' in the supposition "Let *a* be one of the spheres in Black's symmetric world' indeed refers singularly to only one of the two indiscernible spheres in Black's world.

Whether the same idea applies to indiscernible non-individuals will be the subject of Chapter 8. There, I will show that unlike indiscernible individuals, indiscernible non-individuals cannot be referred to by any device of singular reference. This, however, will turn out to be a consequence of their non-individuality, and not a consequence of their indiscernibility.

Chapter 7

Probabilistic Reference

As we saw in Chapter 6, Arbitrary Reference (AR) is the idea that we can refer to individuals with some degree of arbitrariness. Although there are different accounts of Arbitrary Reference, nearly all of them can be challenged on the basis that they entail the existence of free-floating semantic facts, namely: semantic facts which are not grounded in any non-semantic fact.

Here I propose a solution. First, I argue that friends of AR can answer the challenge by appealing to the notion of indeterministic grounding. Then, I propose a new account of Arbitrary Reference as a probabilistic phenomenon, and argue that this new account should be preferred over the classical versions of AR.

This Chapter is divided into seven Sections. In Section 7.1, I briefly discuss three recent accounts of AR developed respectively in Breckenridge & Magidor (2012), Martino (2001), and Woods (2014). In Section 7.2 I introduce the challenge from free-floating semantic facts to the impossibility of Arbitrary Reference. According to this challenge, Arbitrary Reference postulates the existence of semantic facts which are not grounded in any non-semantic fact. However, any semantic fact must be grounded in some non-semantic fact. From this it follows that Arbitrary Reference is impossible. In Section 7.3 I discuss the notion of indeterministic grounding and argue that it allows the friends of AR to overcome the challenge discussed in Section 7.2 without having to commit to the existence of ungrounded semantic facts. In Section 7.4 I present a new theory of AR, according to which Arbitrary Reference is best understood as a probabilistic phenomenon, and argue (Section 7.5) that this new account is superior to the theories of AR currently on the market. Finally, in Section 7.6, I show how this account can help us develop a theory of Arbitrary Reference devoid of free-floating semantic facts.

7.1 Arbitrary Reference

Consider the following proof that every natural number is either even or odd:

> Base case: Let $n = 1$. Since $1 = 2 \times 0 + 1$, then n is odd by definition.
>
> Induction step: Let n be an arbitrary natural number greater than 1, and assume (for induction) that $n - 1$ is either even or odd. Now, if $n - 1$ is even then n is odd, for *it* is the sum of an even number and an odd one. If instead $n - 1$ is odd then n is even, for *it* is the sum of two odd numbers.
>
> By induction, every natural number is either even or odd.

When we say: "Let n be a natural number [...]", is 'n' something like a proper name? If it is, then what does it refer to? If, on the other hand, it is not, then what is it? According to all extant the theories of Arbitrary Reference, 'n' is indeed a proper name. The same goes for the instantial terms 'Pierre' and '$\mathcal{M}$' in sentences like "Let Pierre be an arbitrary French man" and "Let $\mathcal{M}$ be a non-standard model of first order Peano Arithmetic".[1]

The main idea behind AR is that in certain situations we can use proper names to refer to individual entities in an arbitrary way. In this Chapter, I focus on the three accounts of AR already presented in Sections 6.3 and 6.4, all of which agree with the fact that instantial terms refer, with some degrees of arbitrariness, to non arbitrary individuals.[2]

The reader will remember that Breckenridge and Magidor (2012, p. 377) define AR as:

> **Arbitrary Reference (AR)**: It is possible to fix the reference of an expression arbitrarily. When we do so, the expression receives its ordinary kind of semantic value, though we do not and cannot know which value in particular it receives.

According to Breckenridge and Magidor (2012), cases of arbitrary reference are cases where the reference of some linguistic entities is not grounded in any non-semantic fact. To explain. Consider again the mathematical proof mentioned above. According to Breckenridge and Magidor, in the context of the proof, the term 'n' in the supposition "Let n be a natural number" is assigned a particular natural number, say 55. Importantly, there is no fact of

[1]As per King (1991, 239), instantial terms are "expressions of generality": terms which name objects in an indefinite way, like variables.

[2]For the idea that AR involves reference to arbitrary entities instead, see Fine & Tennant (1983), Fine (1985a) and Fine (1985c).

the world, no non-semantic fact, that can be used to explain *why* 'n' comes to refer to 55 instead of, say, 29. This semantic fact is, in Breckenridge and Magidor's terminology, fundamental, or ungrounded.[3]

This fundamentality is also the reason why, according to Breckenridge and Magidor, it is impossible to know the reference of an instantial term in cases of arbitrary reference. According to their view, the expression "Let n be a natural number [...]" is associated to a choice function $f : \wp(\mathbb{N}) \rightarrow \mathbb{N}$ from the power set of the natural numbers to the natural numbers themselves, and the output of this function when its input is $\mathbb{N}$ is the referent of the term 'n'.

A similar view of AR is proposed by Martino (2001), who suggests that we can explain how arbitrary reference works by imagining that we, as speakers, have direct access to some ideal agent, which in turn has direct access to all the individuals in our domain of discourse. Then, whenever we write down the supposition "Let n be a natural number", the ideal agent we have access to chooses a particular number in the set of natural numbers, and names it 'n'. According to Martino (2001), this is what fixes the reference of the instantial term 'n'. Therefore, in Martino's view too, expressions like "Let n be a natural number" are associated with a choice function $f : \wp(\mathbb{N}) \rightarrow \mathbb{N}$, whose output for $\mathbb{N}$ is the referent of 'n'.

Unlike Breckenridge and Magidor (2012) and Martino (2001), Woods sug-

[3]In the literature about Arbitrary Reference, it is common to use the notions of 'grounding', 'dependence' and 'determination' almost interchangeably. Following Audi (2012a) and Rosen (2010), I will here use 'grounding' as a relation of metaphysical dependence between facts. What I will be saying, however, is easily translatable in the language of Cameron (2008) and Schaffer (2009; 2010), who believe grounding is a categorically neutral relation. My account is therefore not committed to the existence of facts, nor to any particular feature that might be deemed essential to facts. I use facts in my exposition, but nothing will hinge on this.

gests a supervaluationist account of Arbitrary Reference, according to which suppositions like "Let n be a natural number' are associated with entire classes of choice functions. In particular, according to Woods, the expression "Let n be a natural number" is associated with the class of total functions $f : \wp(\mathbb{N}) \to \mathbb{N}$ such that to each subset X of $\mathbb{N}$, these functions associate an element x in X. Furthermore, these functions associate to the empty subset one arbitrary element of $\mathbb{N}$.[4]

According to Woods, the term 'n' in the supposition "Let n be a natural number" refers *over* the set of natural numbers, and has distinct unique referents according to distinct precisifications.[5] In particular, each precisification corresponds to one choice function, and delivers a unique referent for 'n'. Finally, Woods holds that, whenever an expression involving arbitrary reference is uttered, there is no determinate fact of the world about which precisification is indeed the one associated with our expression.

7.2 Free-floating semantic facts

All these accounts can be challenged on the basis that they require some semantic facts to be free-floating. In particular, as we will see, none of these views of AR can explain what determines the semantic fact that instantial terms like 'n' come to refer to the particular individual they refer to. We

[4]This last condition ensures that our choice functions will correctly model the semantic behaviour of Hilbert's ε operator. See Section 6.4.2.

[5]Boccuni and Woods (2020, 309) use the locution 'refers *over*' to indicate that instantial terms like 'n' refer to single entities in the relevant domain of quantification *only when* a precisification is selected. When no precisification is selected, these terms do not have single referents: instead, they behave like variables, and range over a class of potential referents.

say that a semantic fact is free-floating when it is not grounded in any non-semantic fact. (Examples of non-semantic facts are facts about the use of an expression or facts concerning the context in which an expression is uttered.)

To better understand the challenge consider again the proof presented in Section 7.1, and let w_1 and w_2 be two possible worlds which agree with respect to all non-semantic facts. Suppose that w_1 is such that Charlie, a first year student at Oxford University, writes down the proof that every natural number is either even or odd. By stipulation, w_2 is also such that Charlie, a first year student at Oxford University, writes down the proof that every natural number is either even or odd.[6] Furthermore, since w_1 and w_2 agree with respect to all non-semantic facts, the context and the use facts associated with Charlie's proof in w_1 must be the same as the context and the use facts associated with Charlie's proof in w_2. The challenge arises because AR allows for the fact that the referent of the term 'n' in w_1 might not be the same as the referent of the term 'n' in w_2 — despite all non-semantic facts are exactly the same in w_1 and w_2.

According to Breckenridge and Magidor (2012), this happens because two distinct choice functions might be associated to the expression "Let n be a natural number" in w_1 and w_2 respectively. Martino's (2001) account is similar: distinct worlds, (possibly) distinct choice functions. Again, in game theoretic terms, we can make sense of this by imagining that the ideal agent Charlie has direct access to in both w_1 and w_2 can preform two distinct choice

[6]Whether the Charlie in w_2 is the same individual as the Charlie in w_1, or just one of his counterparts, is irrelevant for what follows. For this reason, I will here remain silent on the problem of transworld individuals and counterparts.

acts in w_1 and w_2 respectively.[7] According to Woods's (2014) account, the expression "Let n be a natural number" is associated with the *same* class of (total) choice functions in both w_1 and w_2. However, since any possible world correspond to an admissible precisification, it should be easy to see that the referent of the instantial term 'n' w_1 will be different from the referent of 'n' in w_2.[8]

This shift in referents between w_1 and w_2 is to be expected. Recall in fact that the main idea of AR is that, in certain situations, our referential practices have some degree of arbitrariness. And fixing a choice function or a precisification in all possible worlds which agree with respect to all non-semantic facts would make the entire idea of an arbitrary reference collapse.

However, it is a common assumption in the Philosophy of Language that every semantic fact is grounded in some non-semantic facts. We find this assumption in virtually all theories of reference, be them internalist (see Dummett 1993 and Fodor 1987) or externalist (see Putnam 1975 and Kripke 1980). Let's call this the Grounding Principle (GP).

Although this challenge affects all the theories outlined in Section 7.1, only Breckenridge and Magidor (2012) explicitly discuss it in relation to Arbitrary Reference. They first consider the objection that if it is indeed the case that by stipulating "Let n be a natural number" we manage to refer to a particular natural number, then something must determine which number

[7] Again, nothing changes if Charlie is not a transworld individual.

[8] That Woods's account of AR is committed to free-floating semantic facts has nothing to do with the fact that it is a supervaluationist account. Supervaluationist accounts, in fact, are generally not committed to free-floating semantic facts. What generates this committment in the case of Woods's account is that, since the term 'n' is supposed to refer *arbitrarily*, there is no stipulation, or non-semantic fact, which determines which entity is the referent of 'n' in any precisification.

we successfully referred to. To this, they answer that there is nothing in the world that determines which number we have referred to with our stipulation, except for the very fact that we have successfully referred to the number we referred to. They therefore reject the widely held view that all semantic facts are grounded in non-semantic facts.

The thesis that some semantic facts are simply ungrounded is also defended in Kearns & Magidor (2012). Here, the authors consider several challenges to the thesis that every semantic fact supervenes on some non-semantic facts (*Semantic Supervenience*), and reject it in favour of what they call Semantic Sovereignty. Although I doubt that what Kearns and Magidor call *Semantic Supervenience* is equivalent to GP, the important thing to notice here is that one might find the thesis that some semantic facts are fundamental independently plausible, and therefore be satisfied with the theories of AR discussed above, even if they demand for some free-floating semantic facts.

However, I think that a different approach to AR is viable, which makes it consistent with at least some weak form of GP. And since the Grounding Principle seems to be still virtually universally accepted, I find it interesting to try and give a version of AR that is compatible with it.

7.3 Indeterministic Grounding

Before expanding on this new approach to AR I want to argue that an appeal to the notion of *indeterministic grounding*, which seems to have passed under the radars of AR theorists, can help showing that existent accounts of AR are indeed compatible with a weaker version of GP, thereby evading the challenge from free-floating facts. To make my argument precise, it will be useful to characterize the notion of *grounding* in its full extent. What I say here will also apply to the following Sections.

As I have already mentioned in Section 7.1, by grounding I mean a relation of metaphysical dependence between facts. Whether this relation is one of *explanation* or *determination* is orthogonal to the present purposes. Therefore I will remain silent on the issue, as to make the argument appealing to unionists and separatists alike.[9] If P is a proposition, I will write '[P]' for the fact which corresponds to P. Therefore, in what follows, '[Plato was a student of Socrates]' will indicate *the fact that* Plato was a student of Socrates. (Similarly for other propositions.)

In what follows I will not make use of any property which is commonly associated with grounding: i.e. transitivity, asymmetry, irreflexivity, non-monotonicity, and well-foundedness. My arguments will therefore be palatable also for those who believe grounding not to be transitive, asymmetric, irreflexive, non-monotonic, or non-well-founded.[10]

GP is commonly understood as the thesis that every semantic fact is *deterministically grounded* in some non-semantic facts. This has led to the idea that if some semantic facts are not deterministically grounded in non-semantic facts, then there must be some fundamental semantic facts — and

[9]For the idea that grounding should be thought as a relation of explanation, see: Dasgupta (2017) and Fine (2012). For the idea that grounding is instead better understood as a kind of determination, see: Audi (2012b) and Schaffer (2009).

[10]For an extensive treatment of the of the various interesting combinations of formal properties of grounding, see Bliss & Priest (2018). For an interesting defence of the idea that grounding is non-well-founded, see Cameron (2022).

this flies in the face of the strong intuition that semantic facts are never fundamental. However, there is a weaker notion of grounding, that of *indeterministic grounding*, which has been increasingly discussed in the literature about metaphysical dependence, and which can help breaking down the implication from a lack of deterministic grounding to fundamentality. (See, among others, Bader 2021, Craver 2017 and Montero 2013.)

Indeterministic grounding happens when some low-level facts $[P_1]$ ... $[P_n]$, which provide a full grounding base for some incompatible high-level facts $[Q_1]$ and $[Q_2]$, underdetermine which of $[Q_1]$ or $[Q_2]$ obtains. In this case, although $[P_1]$ ... $[P_n]$ fully ground $[Q_1]$ and $[Q_2]$, the relation of grounding between low- and high-level facts is indeterministic: the obtaining of $[P_1]$ ... $[P_n]$ alone doesn't suffice for the obtaining of $[Q_1]$ or $[Q_2]$. Some form of chance is required, which however doesn't threaten the grounding chain: whichever of $[Q_1]$ and $[Q_2]$ obtains, it will be grounded, albeit indeterministically, on $[P_1]$ ... $[P_n]$. As Bader (2021, p. 1123) remarks: "Something that is [indeterministically] grounded is to some extent grounded and to some extent brute. Though its ground is incomplete, *it does have a ground*. Accordingly, it is not fundamental but derivative." (My emphasis.) And: "What lacks a ground is not the range of possible outcomes, but only the way in which chance [...] fixes one of [them]." (Bader 2021, p. 1126)

It is worth stressing that although the introduction of chance/probability in the picture of metaphysical dependence is likely to raise some eyebrows, a probabilistic version of grounding is less problematic than it might look at first sight. Many fundamental laws of our best scientific theories are probabilistic or statistical laws. Quantum Mechanics itself is, for example, a probabilistic theory.

Now if the reality we inhabit is (at least contingently) indeterministic, why should we expect our explanations to be always deterministic? If you believe that Metaphysics should give a faithful representation of reality where connections of metaphysical dependence are transparent, then you should believe that, if reality is indeterministic, then an indeterministic representation of it should be more faithful than a merely deterministic one. Finally, many of our scientific theories predict reality to be indeterministic — and therefore, instead of been looked at with scepticism, the inclusion of probability in our metaphysical theories should indeed be welcomed.[11] (Here I will remain silent on what probabilities are, for the above argument and the account of AR I will propose are compatible with multiple interpretations of probability. The only restriction is that probabilities should not be thought of as subjective.)

How do we apply the notion of *indeterministic grounding* to the accounts of AR discussed in Section 7.1? Breckenridge and Magidor's and Martino's accounts can be dealt with together. In both views, we can explain the relevant grounding between semantic and non-semantic facts as follows. When the sentence "Let n be a natural number" is uttered, some non-semantic facts (including use facts and facts about the context in which the sentence is uttered) determine a set of suitable choice functions that can be associated to the sentence. In this case, any choice function $c : \wp(\mathbb{N}) \to \mathbb{N}$ from the power set of the natural numbers to the natural numbers is a suitable function.[12]

[11] For other arguments along these lines, see again Bader (2021) and Craver (2017).

[12] Notice that not *any* choice function will do. No choice function $c : \wp(\mathbb{Z}^-) \to \mathbb{Z}^-$ from the power set of the negative integers to the negative integers, for instance, is a suitable

Let $[U_1]$ ... $[U_n]$ be the relevant non-semantic facts and $[N]$ be the fact that n is a natural number. Then, although $[U_1]$ ... $[U_n]$ and $[N]$ constitute a full grounding base for whatever choice function gets associated to "Let n be a natural number", their obtaining alone doesn't suffice to determine *which* choice function gets in fact associated to the expression at hand — thereby underdetermining *which* number is the referent of the instantial term 'n'.

Using indeterministic grounding, we obtain the following grounding chain: the semantic fact [the number x is the referent of 'n'] is fully grounded in [c is the function associated to the relevant expression], which is in turn *indeterministically grounded* in $[U_1]$ ... $[U_n]$ and $[N]$, which are non-semantic facts if anything is. Again, following Bader (2021), the relevant semantic fact is indeed grounded in non-semantic facts, and is therefore not fundamental. What lacks a ground is the way in which chance fixes c among all the possible choice functions. However this is hardly a semantic fact: for it is a fact *about* grounding, which is a metaphysical, and not a semantic relation.

The same explanation, with some adjustments, can be used in Woods's case. Recall that according to Woods (2014), the expression "Let n be a natural number" is associated with the whole set X of choice functions from the power set of $\mathbb{N}$ to $\mathbb{N}$, and that the referent of the instantial term 'n' depends, in Woods's account, on the relevant precisification at the context of utterance. Notice that the set X of choice functions that in Woods's account of AR is associated to the expression "Let n be a natural number" is the same as the set of suitable choice functions determined, in the case of Martino's and Breckenridge and Magidor's accounts, by the non-semantic facts $[U_1]$

function for the expression "Let n be a natural number".

... $[U_n]$. Notice further that a precisification, in Woods's account, is nothing more than the singling out of a choice function in the set of suitable functions associated to the relevant arbitrary expression.

With these considerations in place, it should be clear that the relevant grounding relations between semantic and non-semantic facts can be explained, in Woods's case, exactly as they have been explained in Martino's and Breckenridge and Magidor's cases. More precisely we can say that, in Woods's account of AR, when the sentence "Let n be a natural number" is uttered, some non-semantic facts $[U_1]$ ... $[U_n]$ determine the set of suitable choice functions associated to the sentence. Where $[X]$ is the fact that all and only the choice functions in X are suitable for the expression at hand, we say that although $[U_1]$ ... $[U_n]$ and $[X]$ constitute a full grounding base for whatever precisification is relevant for the context of utterance, their obtaining alone doesn't suffice to determine *which* precisification obtains, that is: *which* of the many suitable choice functions gets associated to the relevant expression.[13] So again, we can say that the semantic fact [the number x is the referent of 'n'] is fully grounded in [P is the relevant precisification], which is in turn *indeterministically grounded* in $[U_1]$ ... $[U_n]$ and $[X]$, which are non-semantic facts.[14]

With the notion of indeterministic grounding in their conceptual toolbox, friends of AR can now resist the challenge from free-floating facts in two ways. The first is to argue that a correct interpretation of GP must include indeterministic grounding. If so, GP should be understood as the principle

[13]Notice that, in Woods's account, $[X]$ is equivalent to $[N]$.

[14]Notice again that [P is the relevant precisification] is equivalent, in Woods's account, to [c is the function associated to the relevant expression].

that any semantic fact is grounded, deterministically or indeterministically, in some non-semantic facts. And since it can be shown that the relevant semantic facts in AR are indeed indeterministically grounded in non-semantic facts, then AR is compatible with GP and the challenge is misplaced. A second way is to argue that, although GP should be interpreted only in terms of deterministic grounding, still the notion of indeterministic grounding can be used to show that, although not deterministically grounded in any non-semantic facts, the free-floating facts the AR theorists are challenged with are far from fundamental. And it seems that what lies at the heart of the challenge to AR is that semantic facts should not be fundamental. That is: indeterministic grounding can be used to reassure the opponent of AR that there is indeed no reason to worry, for although a strict version of GP is violated by AR, still the problematic free-floating facts are not in any way fundamental.

7.4 Probabilistic Reference

What would happen if we thought of Arbitrary Reference as a probabilistic phenomenon? The accounts of AR we discussed in Section 7.1 lend themselves quite naturally, I think, to a probabilistic reading.

Informally, the idea is that when we make suppositions like "Let n be a natural number" we introduce some probabilistic constraints on the possible referents of the instantial term 'n'. I believe that any AR theorist would claim that, given the supposition "Let n be a natural number", any natural number *has the same probability as* any other natural number to be the referent of

'n'. (That is: any natural number is as good a candidate as any other natural number to be the referent of 'n'.) The same goes with the supposition "Let Pierre be an arbitrary French man". I believe any AR theorist would agree that any French man has the same probability as any other to be the referent of the term 'Pierre', granted the supposition at hand.

As I see it, these claims are not merely *consistent with* the accounts of AR discussed above; indeed, they seem to me to follow from the very same notion of *arbitrariness* the AR theorists employ. Any account of AR which claimed, say, that certain natural numbers are better candidates then others to be the referents of 'n' given only the supposition "Let n be a natural number" would sound extremely suspicious: for what would it mean, in such case, that 'n' refers *arbitrarily*? This is why I suggest we could look at this notion of *arbitrariness* through the lens of probability. This would give us a new account of Arbitrary Reference, one in which AR is understood as inherently probabilistic. (In the remainder I will call this new account PAR, for Probabilistic Arbitrary Reference.)

To see how such an account might work, consider again the supposition "Let n be a natural number". According to PAR, instead of a choice function or a set of (total) choice functions, the relevant supposition is assigned a probability function $p : \mathbb{N} \to [0, 1]$ from the set of natural numbers to the closed interval [0, 1]. This function assigns to each element x in $\mathbb{N}$ a value which represents the probability that x is the referent of 'n'.

Two clarifications are in order. First, that it seem a plausible assumption that the domain of the relevant probability function will depend on the expression at hand. Take for example the sentence "Let n be a natural number". The probability function associated with it will plausibly have the $\mathbb{N}$ as a domain. Were we to consider the sentence "Let r be a real number", the domain of the relevant probability function would plausibly be the set $\mathbb{R}$ of real numbers. Finally, were we to consider "Let Pierre be an arbitrary French man", the domain of the assigned probability function would plausibly be the set F of French men. (Though nothing really hinges on this: we could stipulate for instance that any probability function had as domain the set NC of all the possible concrete and non-concrete entities, as per Williamson (1998). As long as the assignment of probabilities to the relevant possible referents is correct, the domain of our probability functions will not matter — provided it contains at least the entities in the universe of discourse of the relevant expression.)

Second, we define the co-domain of our probability function as a closed interval in the hyperreal numbers $^*\mathbb{R}$, as per A. Robinson (1961) and A. Robinson (1966). The issue deriving from letting our co-domain be an interval over the real numbers is that, when the function associates the same value to all the elements of an infinite set, the value associated to each element is identified with 0 by definition.[15]

For example, consider again the sentence "Let n be a natural number" and assume, *pace* finitist doubts, that there are infinitely many natural numbers. Now: according to PAR any natural number has the same probability as any other to be the referent of 'n'. This is represented by the fact that the probability function $p : \mathbb{N} \to [0, 1]$ assigned to the expression at hand is

[15] To be precise, we can define p's co-domain as an interval in any system which includes infinitesimals, like the superreals or Conway's numbers. For an interesting approach to infinitesimal probabilities with Conway's numbers, see Chen & Rubio (2020).

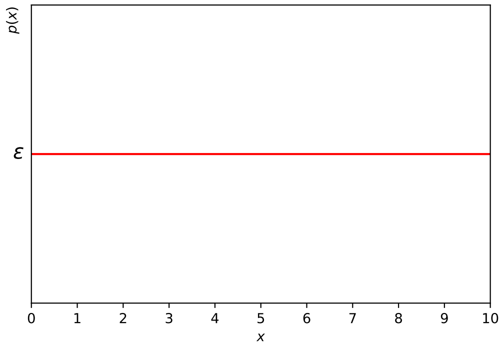
constant, that is: p associates to each member of $\mathbb{N}$ the same value. However, being a probability function, the sum of all $p(x)$ such that $x \in \mathbb{N}$ must be equal to 1. Therefore, were we to assign a *real* value to the members of $\mathbb{N}$, this value couldn't be different from 0. (And this would mean that, for any x in $\mathbb{N}$, the probability that x is the referent of 'n' is null.) This would in turn entail that no natural number could be the referent of 'n', and therefore that if 'n' was a name, it would be an empty one: and this would run contrary to the fundamental intuition behind Arbitrary Reference — that *there is* some natural number which is the referent of the term 'n', even if we don't and cannot know which one it is.

To solve this issue we can define p's co-domain as a subset of the hyperreal numbers. In this way, even in case p is a constant function defined for an infinite domain, the value associated to each entry of p will be strictly greater than 0 — albeit being still strictly smaller than any real number. In the case of the supposition "Let n be a natural number", for any x in $\mathbb{N}$, $p(x) = \varepsilon$, where ε is an infinitesimal quantity.

In this way we can make sense of the intuitions that (1) each natural number has the same probability as any other natural number to be the referent of 'n', and that (2) each natural number has a non null probability to be the referent of 'n'. I suggest that, as in the case of probabilities, the appeal to infinitesimal quantities should not be taken as a reason to worry: infinitesimals are used every day in virtually any branch of physics, and their treatment has been made mathematically and logically rigorous by non-standard analysis and smooth infinitesimal analysis.[16]

[16] A thorough discussion of infinitesimal probabilities would take us too far afield. The

A close-up of the probability function p assigned to the supposition "Let n be a natural number" is the following:



where the value assigned to any number is ε, an infinitesimal quantity.[17] For a more interesting case, one can consider the sentence "Let X be one of the final candidates at the 2016 US presidential elections".

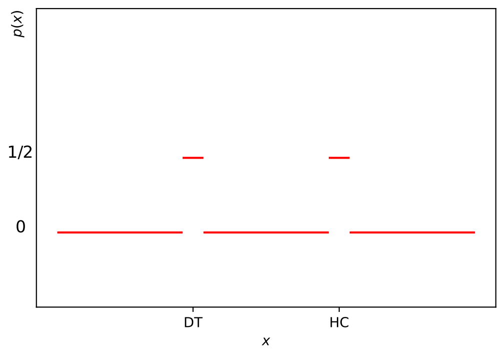
In this case, either 'X' refers to Donald Trump, or 'X' refers to Hillary Clinton. Since no further constraint on reference is specified, PAR holds, in line with the underlying intuition of Arbitrary Reference, that Hillary Clinton has the same probability as Donald Trump to be the referent of 'X'. Therefore, the function p associated to this sentence will be such that $p(x) = 0.5$ if and only if x is either Donald Trump or Hillary Clinton, and $p(x) = 0$ otherwise.

In the case of the supposition "Let X be one of the final candidates at the 2016 US presidential elections", it might be not so straightforward to answer the question about what the domain of p is supposed to be. Maybe

interested reader is invited to consult Wenmackers (2016). A theory of non-Archimedean probability has been provided in Benci et al. (2013).

[17]Notice that this is just a close-up of the function for the first eleven natural numbers. Were we to zoom out, we would see that p stays constant for all natural numbers greater than 10 too.

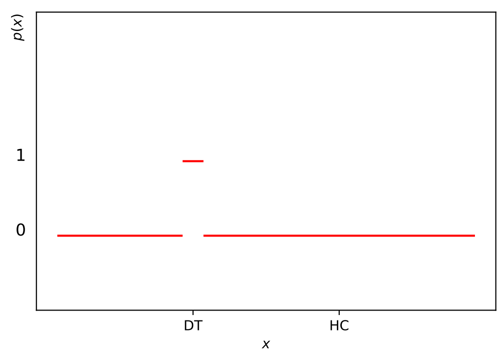
the relevant universe of discourse is the set P of all presidential candidates at the 2016 US elections. Or maybe it is the set C of all US citizens in 2016. Even if I am inclined towards the first of this alternatives, I suggest that the relevant universe of discourse will depend on the context in which the sentence is uttered. What matters for our purposes is that in any of these cases the only non null values will be assigned to Hillary Clinton and Donald Trump. It is also important to notice that since the value associated to the two candidates are both 0.5, then the probability that either Hillary Clinton or Donald Trump is the referent of 'X' is equal to 1, that is: in line with the intuition of AR, 'X' will have a unique referent. The following is a close-up of the probability function p assigned to the supposition "Let X be one of the final candidates at the 2016 US presidential elections":



As you can see, the function is constant apart from two spikes, corresponding to Donald Trump and Hillary Clinton respectively.

7.5 A Comparison

I suggest that PAR is superior to standard AR in numerous respects. First, I suggest that unlike standard AR (except maybe for the account proposed by Woods), PAR can build a bridge between arbitrary reference and canonical non arbitrary reference by considering both of them as instances of the same unified phenomenon. To see how this can be done, it is sufficient to notice that PAR opens the door for a probabilistic understanding of reference *in general*. With PAR, we can argue that *all* referential expressions, and not just those which refer arbitrarily, are associated with probability functions. And if this is the case, then cases of canonical (non arbitrary) reference, like "Donald Trump won the 2016 US presidential elections" are just limit cases of arbitrary reference — cases in which there is only one positive value assigned by the relevant probability function. We can, for example, understand the sentence "Donald Trump won the 2016 US presidential elections" as being associated a probability function p such that $p(x) = 1$ if and only if x is Donald Trump, and $p(x) = 0$ otherwise:



To repeat: with PAR, cases of canonical reference can be defined as those limit cases of arbitrary reference where only one value in the domain of quantification is given a non null value by the probability function associated to the relevant expression. This, I submit, is an advantage of PAR over accounts of AR like the ones of Breckenridge and Magidor, and Martino. It is hard, in fact, to see how these accounts can build a continuity between cases of arbitrary reference and cases of canonical reference.

A second advantage of PAR over standard AR accounts is that it avoids any problems regarding the dependence between distinct choice functions when it comes to expressions containing multiple arbitrarily referring instantial terms. Consider, for instance, the sentence "Let m and n be two natural numbers, such that $n - m = 10$". According to the first two accounts of AR discussed in Section 7.1, this sentence is associated with two choice functions $f : \wp(\mathbb{N}) \to \mathbb{N}$ and $g : \wp(\mathbb{N}) \to \mathbb{N}$ such that $f(\mathbb{N})$ is the referent of 'm' and $g(\mathbb{N})$ is the referent of 'n'.

Notice however, that $f(\mathbb{N})$ and $g(\mathbb{N})$ cannot be just any two natural numbers, for the constraint put on the reference of 'm' and 'n' makes it so that $f(\mathbb{N}) - g(\mathbb{N}) = 10$. But then, f and g cannot just be any two choice functions. They must be related in such a way that the choice of f depends on the choice of g, or *vice versa*. How to explain this dependence is an open question. Breckenridge and Magidor (2012) and Martino (2001) don't consider this question, and neither does Woods (2014) — for whom the challenge must be translated by replacing 'choice functions' with 'sets of choice functions'.

On the contrary, according to PAR, there are not *two* probability functions associated to the relevant expression: there is only *one* function, $p(x, y)$, with two inputs. The constraint imposed on m and n is reflected by the fact that $p(x, y) > 0$ if and only if $y - x = 10$. The graphical representation of this function is a plane embedded in a 3-dimensional space. On the x-axis we have the possible values for x (all natural numbers); on the y-axis we have the possible values for y (again: all natural numbers), and on the z-axis we have the weights relative to the inputs x and y (hyperreal numbers in the closed interval [0, 1]). Visually, we have a flat plane with spikes corresponding to those entries for x and y which meet the constraint specified in the relevant supposition.

The same goes for expressions with more than two arbitrarily referring instantial terms. In general, where φ is a arbitrarily referring expression, n the number of referring instantial terms in φ, and $\psi(\overline{x})$ the constraint imposed on the possible referents: the probabilistic function p associated to φ will be a multivariable functions with n inputs, and will define a n-dimensional hyperplane embedded in a $(n+1)$-dimensional space with spikes corresponding to the list $\overline{x}$ of inputs such that $\psi(\overline{x})$ is true.[18]

This does not only solve the problem of the nature of the dependence between the multiple choice functions associated by standard accounts of AR to expressions containing multiple instantial terms: it also offers a more elegant account of how such reference is achieved. No matter how many instantial terms we have in a given expression, PAR associates to the expression only *one* probability function.

[18] The notation '$\overline{x}$' is a shortcut for a list of variables $x_1, x_2, ... x_n$.

7.6 Grounding and Semantic Vagueness

How does PAR fare against the challenge of free-floating facts? First, I want to notice that PAR is at least as strong as the standard accounts of AR when it comes to meeting this challenge. As with standard AR, we can use the notion of indeterministic grounding to provide a grounding chain for the relevant semantic facts about the reference of instantial terms in arbitrarily referring expressions. Here is how.

Consider again the supposition "Let n be a natural number". According to PAR, when the supposition is made, some non-semantic facts $[U_1]$ ... $[U_n]$, among which is the non-semantic fact $[N]$ *that* n is a natural number, determine a set of weights to be associated with the elements of the relevant domain of quantification (in this case: the natural numbers). All the elements with a weight greater than 0 are suitable referents for the term 'n'. Then, although $[U_1]$ ... $[U_n]$ constitute a full grounding base for whatever element of the domain gets assigned to 'n' as its referent, their obtaining alone doesn't suffice to determine *which* element gets in fact assigned as the referent of 'n'. We can therefore claim that the semantic fact [the number x is the referent of 'n'] is *indeterministically* grounded in [x has a weight greater than 0], which is in turn *deterministically* grounded in $[U_1]$ ... $[U_n]$, which are non-semantic facts by definition.

So the resources available to standard AR to meet the challenge from free-floating semantic facts are also available to PAR. That is: when it comes to meeting this challenge, PAR is at least as strong as AR. However, I believe we can argue that PAR is stronger than standard AR in that it doesn't require the notion of indeterministic grounding to meet the intuition that every semantic fact is ultimately fully grounded in some non-semantic facts. In other words, I believe one could argue that, unlike standard AR, PAR is compatible with the strong version of GP, according to which every semantic fact is deterministically grounded in some non-semantic fact.

Here is how. We could argue that when, in the actual world, one utters the expression "Let n be a natural number", the probability function associated with it determines a distribution of values over all the possible worlds in which the same person uttered the same expression in the same, or in some relevantly similar, context. This distribution assigns a natural number to any such worlds, and, for any world w, the number x assigned to w is the referent of the term 'n' in w. That all these worlds exist is granted by the fact that any number *could have been* the referent of 'n'.[19]

In a spirit sufficiently similar to that of frequentist approaches to probability, we could then define the probability of any number x to be the referent of 'n' as the relative frequency of occurrences of x within the distribution determined by the probability function associated to the relevant expression. With this picture in mind, we can claim that the semantic fact [x is the referent of 'n' in world w] is fully grounded in the non-semantic fact [according to the relevant distribution, world w is assigned value x], and this second fact is fully grounded in the non-semantic fact [the probability function p

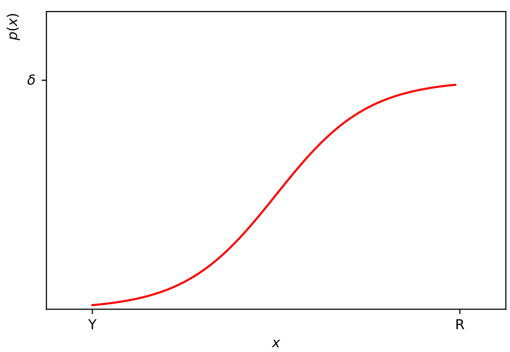
[19] I am here granting the existence of possible worlds, insofar as they are a useful device for the philosophical analysis of modal notions. (See Cowling 2011). It is worth stressing, however, that my argument doesn't depend on any particular claim about their nature. Those sceptical with respect to the notion of 'possible world' can substitute it with that of 'possibility'. Those sceptical with respect to possibilities will not like my story, but so it must be (pun intended).

assigned to the relevant arbitrary expression is such and so]. I suggest that the distribution is *deterministically* grounded in the probability function because the only way I can make sense of the claim that "the same probability function *could have determined* another distribution" is by understanding it as the claim that "if it is true that the natural number x is, at the actual world, the referent of 'n', it is also true that any another number could have been the referent of 'n' at the actual world". But then, isn't this claim made true by the distribution itself?

If the above reasoning is correct, we can argue that PAR fares better than standard AR against the challenge from free-floating semantic facts, for it is compatible with a stronger version of GP. However, I recognise that there are a number of ways to resist the above argument: maybe you think my story about a unique distribution over possible worlds is untenable, or you hold that indeed, in a way I cannot really make sense of, the distribution determined by the probability function could indeed have been different, and therefore that my story shows that the probability function only indeterministically ground the relevant distribution. Anyway, the upshot is that, when it comes to the challenge from ungrounded semantic facts, PAR at least as strong as AR.

There is one last advantage of PAR over AR, which involves the formal treatment of cases of semantic vagueness. In this last paragraph, I want to suggest that, with PAR, we can define semantic vagueness in a new, philosophically interesting way. You might have noticed that the probability functions discussed in Sections 7.4 and 7.5 were rather simple: either they were constant functions or, in cases were the relevant domain of quantification was large enough, they were discontinuous functions with sudden spikes in probability values. This holds for virtually all the common case studies considered in the literature about Arbitrary Reference: when looking at them through the lens of probabilistic reference, we notice that the probability functions associated with them are either constant or discontinuous. This means that either their derivative is null, or they are not differentiable.

There is only one case where this trend breaks: when we consider propositions involving vague concepts. Suppose you have a very long series of colored patches which shade very gradually from yellow to red. Now consider the supposition "Let x be an arbitrary red patch". We expect that, were the series of patch continuous, the probability function associated with the relevant supposition would look as follows:



where δ is some value greater than 0 and such that the area under the function equals 1.[20] If we consider other cases of arbitrary suppositions involving vague concepts the result is similar. What I want to suggest is that, looking at reference in a probabilistic way, we can define semantic vagueness in a

[20] Maybe the actual curve of the function is different: however, this is beside the point I want to make in this paragraph.

new way, by means of the mathematical characteristics of the probability functions associated to vague expressions. In the particular case of PAR, we can define semantic vagueness as the semantic phenomenon that corresponds to differentiable probability functions with non null derivative, when the relevant interval is continuous.

7.7 Conclusion

In this Chapter, I have argued that standard accounts of AR can successfully overcome the challenge from free-floating semantic facts by appealing to the notion of indeterministic grounding. Furthermore, I have suggested a new probabilistic account of Arbitrary Reference, and argued that it has many advantages over classical AR theories: it can build a bridge between arbitrary reference and canonical non arbitrary reference, it offers a more elegant account of how reference is achieved when multiple instantial terms are present in a given supposition, and has the resources to define semantic vagueness in a new, philosophically interesting way.

Chapter 8

Non-individuals and Names

In Chapter 6 I argued that there is a substantial metaphysical distinction between individuals and non-individuals, which is too often overlooked in the debate about the possibility of singular reference to indiscernible entities. I suggested that forgetting this distinction has led, and continues to lead, to two important mistakes.

On the one hand, it has led to assume that all the questions that can meaningfully be asked about indiscernible individuals are equally meaningful when asked about indiscernible non-individuals. On the other hand, it has led to assume that when indeed there are questions which are meaningful for both indiscernible individuals and non-individuals, their answers is independent from the *kind* of indiscernibles at hand.

In this last Chapter I will show that, at least when it comes to the possibility of singular reference, there is no unified answer which applies to indiscernibles across the board. I will do this by showing that singular reference to indiscernible non-individuals is indeed impossible.

When joined together with the main result of Chapter 6, that singular reference to indiscernible individuals is possible, the result in this Chapter will highlight the extent of the confusion which forgetting the distinction between individuals and non-individuals can potentially yield.

I believe the arguments in this Chapter are interesting in and of themselves. For although intuitions abound according to which singular reference to non-individuals is impossible, I am aware of no philosophical argument for why this is indeed the case.

The Chapter is divided into three short Sections. In Section 8.1 I briefly go over the distinction between individuals and non-individuals. In Section 8.2 I discuss the main intuitions about singular reference and non-individuals in the literature, and suggest three distinct arguments for the impossibility of singular reference to non-individuals. These show that no matter whether reference is a function, a non-functional relation, or whether names are just properties, it is impossible to singularly refer to non-individuals. In Section 8.3 I present one further argument against the possibility of singular reference to non-individuals, which doesn't stem from any particular account of reference, and therefore applies across the board.

8.1 Introduction: Non-individuals

An entity x is an 'individual' if and only if the relation of identity applies to x, that is: if and only if x is self-identical, and sentences like "x is distinct from y", where y is an individual too, are meaningful. Contrary to individuals, 'non-individuals' are entities to which the relation of identity doesn't apply. Although we experience our everyday world as a world of individuals, in such a way that it might seem highly counterintuitive to claim that there are entities in this very world for which it doesn't make any sense to say that they are self-identical, many authors have argued that quantum particles in entangled states are non-individuals, in the sense specified above. (See, among others, French 1989.) Williamson (2022, p. 17–18) too seems to consider the possibility of non-individuals. He writes:

> What reason have we to assume that reality does not contain *elusive objects*, incapable in principle of being individually thought of? Although we can think of them collectively — for example, as elusive objects — that is not to single out any one of them in thought. Can we be sure that ordinary material objects do not consist of clouds of elusive sub-sub-atomic particles? We might know them by their collective effects while unable to think of any single one of them. The general question whether there can be elusive objects looks like a good candidate for philosophical consideration.

Now, Williamson (2022) may or may not believe that there are non-individuals. However, if indeed there are non-individuals, his point seems to apply to them, and we should understand non-individuals as elusive objects.

Non-individuals are usually characterised as follows. If x is a non-individual, then sentences like "x is self-identical" and "x is not identical to itself" are meaningless. By the same token, if x and y are non-individuals, then sentences like "x is identical to y" and "x is distinct from y" are equally meaningless. The overall idea is that identity is not universal: although many things are subject to its rule, some things have somehow managed to prosper outside the borders of its realm. (See, among others, Dalla Chiara & Toraldo Di Francia 1993, 1995, Domenech & Holik 2007 French & Krause 2006, and Krause & Coelho 2005.)

8.2 Reference

The existence of non-individuals brings about numerous questions. One of them is whether we can refer to non-individuals in a singular way, as we usually do with individual entities. In what follows I argue that singular reference to non-individuals is impossible. I provide four arguments for this conclusion. The first three arguments address specific accounts of reference, respectively: reference as a function, reference as a non-functional relation, and reference in presence of names which are in turn understood as predicates. The last argument, which is independent from any assumptions about the logical properties of reference, is supposed to be general enough to apply to accounts of reference across the board. With this, I want to fill a gap in the literature about non-individuals: for although many authors have suggested that singular reference to non-individual entities is impossible, no one has as so far discussed why this is the case. (See, among others, Assadian 2019.)

It is interesting to note at the outset that many authors which consider subatomic particles as non-individuals and believe that reference to non-individuals is impossible usually refer either to Schrödinger's (1963) observation that one cannot "mark an electron", or to Weyl's (1950) suggestion that one cannot "demand an alibi of an electron". However, these considerations seem to be about epistemic individuation (as opposed to metaphysical individuation), and most certainly cannot be suggested as reasons for why we cannot singularly refer to non-individuals.[1] Similar considerations apply in fact to indiscernible individuals, too: in his *The Identity of Indiscernibles*, Black (1952, p. 157) explicitly says that we cannot "put a red mark" on either of his spheres, without inevitably changing the nature of his scenario. However, as we saw in Chapter 6, singular reference to Black's spheres, and to indiscernible individuals in general, is indeed possible.

8.2.1 Functional Reference

Suppose that reference is a functional relation, in line with Bach (1987) and Kripke (1980). This means that one and the same name cannot refer to two distinct entities. Now suppose for *reductio* that there is a non-individual, x, such that it is possible to refer to x via a functional reference relation. If this is the case, then there is a name 'n' and a function f such that $f(n) = x$. (That is: if x can be the referent of a name and reference is understood as functional, then there is a reference function f such that x is the output of f when x's name is f's input.) But then, since $f(n)$ is x, we have, by *substitutivity*, that $x = x$. This is in contradiction with the fact that x is a non-individual. I conclude that, if x is a non-individual and reference is functional, singular reference to x is impossible.

A more general result of this argument, which will also be useful later on,

[1] For more on the notions of epistemic and metaphysical individuation, see Lowe 2003. See also Chapter 6 for an argument to the extent that singular reference is possible even in the absence of epistemic individuation.

is that if x is a non-individual, then there is no function f such that x is the output of f for some input. While in my argument I have only mentioned unary functions (i.e. functions with single inputs), the argument can be easily generalised to n-ary functions, for any n. This is particularly relevant if one wants to account for the possibility of other parameters appearing alongside names as inputs of the relevant reference functions (i.e. parameters relative to the context in which the referential expression was uttered).

8.2.2 Relational Reference

Some authors hold that reference is a non-functional relation, and therefore that the same name can simultaneously be associated to more than one referent. (See, among others, Devitt 2015 and Delgado 2019.) To see why singular reference to non-individuals is impossible even with this relaxed notion of reference, suppose that reference is indeed a relation, and that it is not functional. Suppose further (for *reductio*) that there is a non-individual, x, such that it is possible to refer to x via a non-functional reference relation. It follows that there is a relation R and a name 'n' such that $\langle n, x\rangle$ is in R. (R is our reference relation.) But then, we can define a set $A = \{\langle y, z\rangle : \langle y, z\rangle$ is in R and $\langle y, z\rangle = \langle n, x\rangle\}$. By definition, $A = \{\langle n, x\rangle\}$, which means that A is a function which outputs x when given n as input. However, we already saw that x, being a non-individual, cannot be the output of any function. Therefore A cannot exist — and this is a contradiction. I conclude that, if x is a non-individual and reference is a non-functional relation, singular reference to x is impossible.

Again, it is straightforward to see how the argument can be generalised to n-ary relations. Suppose reference is a n-ary relation and x is such that it is possible to refer to it via such relation. Then, there is a relation R and a name 'n' such that $\langle n, t^1, ..., t^{n-2}, x\rangle$ is in R. Then, we can define the set $A = \{\langle n, t^1, ..., t^{n-2}, x\rangle\}$ which is a function which outputs x when given $n, t^1, ..., t^{n-2}$ as inputs. This argument shows, more generally, that if x is a non-individual, then x cannot be the last coordinate of any relation.[2]

8.2.3 Names as Predicates

According to Predicativism names are predicates, and are therefore interpreted, model-theoretically, as subsets of the relevant domain of quantification. (See Burge 1974, Fara 2015, and Gray 2015.) One of the advantages of predicativism over other theories of names is the ability to take natural language expressions like "There are only five Alfreds in the world" at face value. (Notice the pluralisation of the proper name 'Alfred'.)

It is an interesting question whether, within a predicativist framework, non-individuals can be understood as capable of having names.[3] It is even more interesting to realise that we do not need to answer it in order to show that even within a predicativist framework, singular reference to non-individuals is impossible.

This is because, according to the predicativist, reference to singular entities is achieved either via the use of *that*-expressions, like the expression

[2] Provided relations are correctly modeled as sets of n-tuples, and whenever x exists, also its singleton $\{x\}$ exists.

[3] An answer to this question will depend, again, on whether we understand our intended 'model of the world' to be such that, whenever some entity x exists, then its singleton $\{x\}$ exists too.

"That (very) Alfred is tall", or via the use of *the*-expressions, like the expression "The (relevant) Alfred is tall". The semantic behavior of these kinds of expressions, however, is usually understood as functional, whereby the relevant *that*- or *the*-expression is associated to a function which outputs a single entity in the relevant domain of quantification upon receiving the expression (and plausibly some other parameters relative to the context in which the expression was uttered) as inputs. Therefore, the predicativist case collapses into the case, considered above, of functional reference. Again, since a non-individual cannot be the output of any function, singular reference to non-individuals is impossible even within the predicativist framework.

8.3 Lagadonian Languages

Taken at face value, the arguments outlined in Section 8.2 show that singular reference to non-individual entities is impossible. The careful reader, however, will have noticed that a common assumption of all the arguments I have presented is that reference is either a function or a relation, and that, as such, it is best understood as a set of ordered n-tuples. A quick look at the relevant literature reveals that this is in fact the usual assumption: when it comes to formalising relations such as reference, ZFC is virtually universally taken as the background theory.

An alternative way to understand the arguments in Section 8.2 is to take them as witnesses of the fact that non-individuals demand a different understanding of relations and functions. In other words: if one has the intuition that it should be possible, at least in principle, to refer to non-individual entities, then one could conclude, from the arguments in Section 8.2, that ZFC does not characterise reference correctly.

There are some alternatives to ZFC which have been developed explicitly to deal with non-individuals. (See, among others, Dalla Chiara & Toraldo Di Francia 1993 and French & Krause 2006.) Using any of these theories, one could try to characterise reference as a quasi-function or a quasi-relation. Consider, for example French and Krause's Quasi-Set Theory, as developed in Krause (1992), Krause (2004), and French & Krause (2006). Suppose x is a non-individual and and there is a name 'n' and a quasi-function f such that $\langle n, x\rangle$ is in f. By definition, $\langle n, x\rangle = [n, x]$ and $[n, x] = [[n], [n, x]]$.[4] If names are not non-individuals, then $[n, x] = \{n\} \cup [x]$. However, the quasi-set $[x]$ is by definition such that it contains everything which is indiscernible with respect to x. So we might not (and indeed, in most cases we will not) have a unique output for f. Therefore, even though this might be used to plurally refer over a plurality of (indiscernible) non-individuals, it is unclear how any quasi-function can be used as a device of *singular* reference.

Furthermore, I believe there are reasons to be sceptical about the tenability of the axioms of Quasi-Set Theory. One of them is that although the notion of identity is said not to apply to non-individuals, there are other notions, for example indiscernibility, which do apply to them. It is an axiom of the theory, for instance, that if x is a non-individual, then '$x \equiv x$' is true of x — which means that x is indiscernible with respect to itself. But how can this formula be true, if x lacks identity? How can we express this concept,

[4] The notation '$[x, y, z]$' indicates the quasi-set of x, y, and z: e.i. the set of entities which are indiscernible from x, y or z.

if even granted that the variable 'x' could refer to some non-individual, it wouldn't still make sense to say that the first 'x' in the formula refers to the same entity the second 'x' refers too? That is: how can we make sense of the claim that a non-individual is indiscernible with respect to *itself*? Identity here seems presupposed at the outset. (A thorough discussion of Quasi-Set Theory would take us too far afield. For our purposes, these few remarks will suffice.)

However, even if the issues arising from considering reference as a quasi-function are set aside, I believe any such weakened understanding of reference would still be insufficient to guarantee singular reference to non-individual entities. To explain why, I will present one last argument, which aims to show that singular reference to non-individuals is impossible, and that this is so independently of how one understands reference at a formal level.

The main idea behind this argument is that the impossibility of singular reference to non-individual entities is a consequence of what I call the Lagadonian Principle, according to which:

> Some entity x can be singularly referred to in some language $\mathcal{L}$ only if it can be *in principle* singularly referred to in a Lagadonian language.

Lagadonian languages are languages in which each object is its own name, each property is its own predicate, and each relation is its own relation symbol. (See, among others, Roy 1995, p. 219.) Although one usually sees mentions of Lagadonian languages in connection with accounts of possible worlds as maximally consistent sets of sentences,[5] the usefulness of Lagadonian languages extends well beyond the Metaphysics of Modality. One example is Van Bendegem (1999), where a non-standard Lagadonian language is used to argue in favour of strict finitism in the Philosophy of Mathematics: the view that the are only finitely many mathematical entities.[6]

I hold that the Lagadonian Principle is straightforward, for is a direct consequence of the fact that Lagadonian languages afford the weakest possible reference relation: identity. If the Lagadonian Principle is true, then in order to show that singular reference to non-individuals is impossible we just have to show that singular reference to non-individual entities is in principle impossible in a Lagadonian language. But that is easy, for in a Lagadonian language reference collapses into identity: in a Lagadonian language, x refers to y if and only if $x = y$. So suppose that a non-individual entity, x, can be singularly referred to in a Lagadonian language. It follows that there is some non-individual y such that y stands in the identity relation to x. However this is impossible, for since x is a non-individual, x does not stand in the identity relation to any non-individual y. Therefore, if x is a non-individual, there is no Lagadonian language in which x can be singularly referred to. And this entails that, as a consequence of the Lagadonian Principle, there is no language in which x can be singularly referred to.

[5]See, among others: Berto & Jago (2013, p. 87), Kment (2014, ch. 5), Lewis (1986, p. 145–146), and Sider (2002).

[6]On the assumption that reality is infinite, Lagadonian languages are not "learnable languages" in the sense of Davidson (1965), for they will be infinite too. (For discussion, see: Haack 1978.) Relatedly, Read (1997, p. 82–83) notes that the 'downward' Löwenheim-Skolem theorem fails for Lagadonian languages in an uncountably infinite reality. Although not learnable, I hold that Lagadonian languages are still useful theoretical idealisations.

8.4 Conclusion

I believe the argument presented in Section 8.3 does not rely on any particular understanding of reference as a relation or function within any specific formal theory. I take this to show, in connection with the arguments suggested in Section 8.2, that singular reference to non-individual entities is indeed impossible.

This Chapter does not only vindicate the common intuitions one finds in the literature about reference and non-individuals. More importantly, it affords philosophically compelling arguments for the truth of these intuitions. And this, I suggest, is significant. We all know, in fact, that intuitions aren't always a good guide to what is true. And when it comes to entities which are so unconventional, like non-individuals, it is very unclear whether we should rely on intuitions at all.

Further Developments

In this Thesis, I have presented some new results connected with the notion of *indiscernibility* as it is usually defined and used in Metaphysics. Three of these results concern Leibniz's principle of the Identity of Indiscernibles (PII), according to which if entities x and y are qualitatively indiscernible, then x and y are identical. PII connects the notion of indiscernibility to that of numerical identity, by stating that qualitative indiscernibility (i.e. indiscernibility with respect to the class of qualitative properties) is a sufficient condition for individual-identity.

I have argued in Chapter 2 that if PII is understood as affording an analysis of individual-identity, then the weakest acceptable version of PII is one in which the second order quantification is restricted to qualitative properties only. To this conclusion, I have advanced an argument to the extent that any version of PII which quantifies over at least some non-qualitative properties either leaves room for the existence of entities which differ only numerically, or generates an infinite regress in the explanation of the identity of at least some entities.

I have then argued in Chapter 3 that branching worlds (i.e. inherently indeterministic worlds with multiple incompatible time-lines) can be used to construct very strong counterexamples to PII. I have presented one such counterexample, which I called the Disintegrating World, and shown that unlike all other counterexamples to PII discussed in the literatures, the Disintegrating World is successful against all the most common lines of defense of PII.

Finally, I have argued in Chapter 4 that a popular version of PII restricted to ordinary spatio-temporal entities (PII-O) is connected with Haecceitism (the thesis that there are distinct maximal possibilities which include all the same qualitative possibilities) in the following way: if PII-O is not necessarily true, then Haecceitism is true. This result is important for many reasons. First, it shows that the common intuition according to which PII and Haecceitism are independent theses is misguided. Second, it shows that there is a connection between the number of ordinary spatio-temporal individuals and the number of overall possibilities. Finally, it represents a strong argument in favour of Haecceitism, for it entails that the necessary truth of PII is a consequence of Anti-Haecceitism.

In Chapter 5 I focused on qualitative and non-qualitative properties. This is because any definition of the notion of *qualitative indiscernibility*, which plays a major role in philosophical debates about indiscernibles, inevitably includes a definition of qualitative and non-qualitative properties. I have proposed a new account of qualitative properties, according to which a property P is qualitative if and only if P's distribution is invariant under possible and impossible identity assignments.

In Chapters 6, 7 and 8 I looked at indiscernibles from the point of view of the Philosophy of Language, and attempted to answer the question whether it is possible to singularly refer to only one among many indiscernible entities. Contrary to the intuitions in the literature, I argued in Chapter 6 that there is no compelling reason to believe that one cannot singularly refer to indiscernibles, provided the indiscernibles at hand are individuals (i.e. entities to which identity applies), and the relevant reference-fixing mechanism contains some element of arbitrariness. To this end, I discussed some accounts of Arbitrary Reference and showed that they are compatible with the possibility of singular reference to indiscernible individuals.

However, since Arbitrary Reference is usually challenged on the basis that it seems to demand that at least some semantic facts are fundamental, I devoted the first half of Chapter 7 to a defense of Arbitrary Reference. In particular, I argued that Arbitrary Reference entails the existence of fundamental semantic facts only if the relevant notion of *grounding* is deemed to be that of *deterministic grounding*. If in fact one opts for a more relaxed notion of grounding, like the notion of *indeterministic grounding*, then the entailment from Arbitrary Reference to fundamental semantic facts breaks. In the second half of Chapter 7 I have advanced a new account of Arbitrary Reference according to which Arbitrary Reference is an inherently probabilistic phenomenon. I have discussed the details of my new proposal, and argued that it fares better then the extant accounts of Arbitrary Reference in multiple respects.

I concluded my Thesis with Chapter 8, in which I have put forward four arguments for the conclusion that, unlike singular reference to indiscernible individuals, singular reference to indiscernible non-individuals (i.e. entities to which identity does *not* apply) is impossible. With this, I wanted to fill a gap in the literature about non-individuals. Although in fact the intuition that singular reference to non-individuals is impossible is widespread in the literature, no argument has been suggested to date for why this is the case.

I believe that there are many paths in which the work carried out in this Thesis can be further developed. I will here outline two, which are the ones I find myself most interested in. The first one has to do with the interplay of the notions of *indiscernibility*, *individuality*, and *parthood*. For suppose for a moment that the Received View of Quantum Mechanics, as discussed in Chapter 4 and Chapter 8, is correct. Then, elementary particles in entangled states are not only indiscernible, but also non-inviduals. However, we know that these particles make up all the spatio-temporal entities we interact with in our daily experience — and these entities are not only discernible: they are also individuals. Then: How is it possible that mereological fusions of indiscernible non-individuals result in discernible individuals? What does this tell us about the notions of *indiscernibility* and *individuality*? Furthermore: Is it possible to construct an extended mereological theory which can describe this phenomenon? And is there a precise point where *individuality* and *indiscernibility* emerge?

The second development I am interested in concerns how we represent indiscernible entities in our models. There are two remarks in the literature about indiscernibles which are of particular interest to me. The first one is that Zermelo-Fraenkel's Set Theory with Choice (ZFC) is not a good ambient theory for models which contain indiscernibles. The second one is that this is due to the Axiom of Extensionality, which makes any two distinct sets discernible. Although I believe the first claim is true, since it is true that in ZFC every two distinct sets are discernible, I believe the second claim is false. This is because the Axiom of Extensionality features also in Boffa Set Theory — but it is not true that any two distinct Boffa sets are discernible from within the theory. (For reasons of space, I will not expand on why this is the case. The interested reader is referred to Aczel 1988 and Rieger 2000.) But then, there must be something else which makes ZFC unsuitable as the background theory for models with indiscernibles. And maybe, instead of resorting to *urelemente*, as in French & Krause (2006), we could use Boffa sets as models for indiscernibles.

I thank anyone who has arrived to this last page. I hope you found this worthwhile.

References

Aczel, P. (1988). *Non-Well-Founded Sets (CSLI Lecture Notes: Number 14)*. Stanford: CSLI Publications.

Adams, R. M. (1979). Primitive Thisness and Primitive Identity. *Journal of Philosophy*, *76*(1), 5–26. doi: 10.2307/2025812

Adams, R. M. (1981). Actualism and Thisness. *Synthese*, *49*(1), 3–41. doi: 10.1007/bf01063914

Allen, S. R. (2016). *A Critical Introduction to Properties*. London, UK: Bloomsbury.

Ariew, R., & Garber, D. (1989). *G. W. Leibniz Philosophical Essays*. Hackett.

Armstrong, D. M. (1978). *Universals and Scientific Realism: Nominalism and Realism Vol. I*. Cambridge University Press.

Armstrong, D. M. (1980). Identity Through Time. In P. Van Inwagen (Ed.), *Time and Cause: Essays Presented to Richard Taylor* (pp. 67–78). Reidel.

Armstrong, D. M. (1989). *A Combinatorial Theory of Possibility*. Cambridge and New York: Cambridge University Press.

Assadian, B. (2018). The Semantic Plights of the Ante-Rem Structuralist. *Philosophical Studies*, *175*(12), 1–20. doi: 10.1007/s11098-017-1001-7

Assadian, B. (2019). In Defence of Utterly Indiscernible Entities. *Philosophical Studies*, *176*(10), 2551–2561. doi: 10.1007/s11098-018-1140-5

Audi, P. (2012a). A Clarification and Defense of the Notion of Grounding. In F. Correia & B. Schnieder (Eds.), *Metaphysical Grounding: Understanding the Structure of Reality* (pp. 101–121). Cambridge University Press.

Audi, P. (2012b). Grounding: Toward a Theory of the "In-Virtue-Of" Relation. *Journal of Philosophy*, *109*(12), 685–711. doi: 10.5840/jphil20121091232

Ayer, A. J. (1954). *Philosophical Essays.* Westport, Conn.: Greenwood Press.

Bach, K. (1987). *Thought and Reference.* New York: Oxford University Press.

Bader, R. (2021). The Fundamental and the Brute. *Philosophical Studies*, *178*(4), 1121–1142. doi: 10.1007/s11098-020-01486-z

Barcan, R. C. (1947). The Identity of Individuals in a Strict Functional Calculus of Second Order. *Journal of Symbolic Logic*, *12*(1), 12–15. doi: 10.2307/2267171

Barnes, E., & Cameron, R. (2011). Back to the Open Future. *Philosophical Perspectives*, *25*(1), 1–26. doi: 10.1111/j.1520-8583.2011.00206.x

Belnap, N. (1992). Branching Space-Time. *Synthese*, *92*(3), 385–434. doi: 10.1007/bf00414289

Belnap, N., & Green, M. (1994). Indeterminism and the Thin Red Line. *Philosophical Perspectives*, *8*, 365–388. doi: 10.2307/2214178

Benci, V., Horsten, L., & Wenmackers, S. (2013). Non-Archimedean Probability. *Milan Journal of Mathematics*, *81*(1), 121–151.

Bennett, K. (2004). Global Supervenience and Dependence. *Philosophy and Phenomenological Research*, *68*(3), 501–529. doi: 10.1111/j.1933-1592.2004.tb00364.x

Berto, F. (2017). Counting the Particles: Entity and Identity in the Philosophy of Physics. *Metaphysica*, *18*(1). doi: 10.1515/mp-2017-0001

Berto, F., & Jago, M. (2013). *Impossible Worlds*. Oxford: Oxford University Press.

Black, M. (1952). The Identity of Indiscernibles. *Mind*, *61*(242), 153–164. doi: 10.1093/mind/lxi.242.153

Bliss, R., & Priest, G. (2018). *Reality and its Structure: Essays in Fundamentality*. Oxford, UK: Oxford University Press.

Boccuni, F., & Woods, J. (2020). Structuralist Neologicism? *Philosophia Mathematica*, *28*(3), 296–316. doi: 10.1093/philmat/nky017

Boffa, M. (1969). Sur la Théorie des Ensembles sans Axiome de Fondement. *Bulletin de la Société Mathématique de Belgique*, *31*, 16–56.

Bolton, M. B. (2021). Kant's Amphiboly as Critique of Leibniz. In *Leibniz and Kant.* Oxford University Press. doi: 10.1093/oso/9780199606368.003.0008

Breckenridge, W., & Magidor, O. (2012). Arbitrary Reference. *Philosophical Studies*, *158*(3), 377–400. doi: 10.1007/s11098-010-9676-z

Brighouse, C. (1997). Determinism and Modality. *British Journal for the Philosophy of Science*, *48*(4), 465–481. doi: 10.1093/bjps/48.4.465

Brook, A., & McRobert, J. (1998). Kant's Attack on the Amphiboly of the Concepts of Reflection. *The Paideia Archive: Twentieth World Congress of Philosophy*, *45*, 41–46. doi: 10.5840/wcp20-paideia199845862

Bueno, O. (2014). Why Identity is Fundamental. *American Philosophical Quarterly*, *51*(4), 325–332.

Burge, T. (1974). Demonstrative Constructions, Reference, and Truth. *Journal of Philosophy*, *71*(7), 205–223. doi: 10.2307/2025348

Burgess, J. P. (1999). Book Review: Stewart Shapiro. Philosophy of Mathematics: Structure and Ontology. *Notre Dame Journal of Formal Logic*, *40*(2), 283–291. doi: 10.1305/ndjfl/1038949543

Button, T. (2006). Realistic Structuralism's Identity Crisis: A Hybrid Solution. *Analysis*, *66*(3), 216–222. doi: 10.1111/j.1467-8284.2006.00617.x

Button, T. (2017). Grades of Discrimination: Indiscernibility, Symmetry, and Relativity. *Notre Dame Journal of Formal Logic*, *58*(4), 527–553. doi: 10.1215/00294527-2017-0007

Cameron, R. P. (2008). Turtles All the Way Down: Regress, Priority and Fundamentality. *Philosophical Quarterly*, *58*(230), 1–14. doi: 10.1111/j.1467-9213.2007.509.x

Cameron, R. P. (2009). Intrinsic and Extrinsic Properties. In R. L. Poidevin (Ed.), *The Routledge Companion to Metaphysics.* Routledge.

Cameron, R. P. (2022). *Chains of Being: Infinite Regress, Circularity, and Metaphysical Explanation.* Oxford: Oxford University Press.

Carnap, R. (1937). *The Logical Syntax of Language.* London: Routledge.

Cartwright, R. (1975). Scattered Objects. In K. Lehrer (Ed.), *Analysis and Metaphysics* (pp. 153–171). Reidel.

Caulton, A. (2013). Discerning "Indistinguishable" Quantum Systems. *Philosophy of Science*, *80*(1), 49–72. doi: 10.1086/668874

Caulton, A., & Butterfield, J. (2012). On Kinds of Indiscernibility in Logic and Metaphysics. *British Journal for the Philosophy of Science*, *63*(1), 27–84. doi: 10.1093/bjps/axr007

Chalmers, D. J. (1996). *The Conscious Mind: In Search of a Fundamental Theory (2nd edition).* Oxford University Press.

Chen, E. K., & Rubio, D. (2020). Surreal Decisions. *Philosophy and Phenomenological Research*, *100*(1), 54–74. doi: 10.1111/phpr.12510

Chisholm, R. M. (1976). *Person and Object: A Metaphysical Study.* London: Open Court.

Chisholm, R. M. (1984). *On Metaphysics*. Univ of Minnesota Press.

Cowling, S. (2011). The Limits of Modality. *Philosophical Quarterly*, *61*(244), 473–495. doi: 10.1111/j.1467-9213.2011.694.x

Cowling, S. (2012). Haecceitism for Modal Realists. *Erkenntnis*, *77*(3), 399–417. doi: 10.1007/s10670-012-9394-3

Cowling, S. (2015). Non-Qualitative Properties. *Erkenntnis*, *80*(2), 275–301. doi: 10.1007/s10670-014-9626-9

Cowling, S. (2017). Conceivability Arguments for Haecceitism. *Synthese*, *194*(10), 4171–4190. doi: 10.1007/s11229-016-1136-7

Cowling, S. (2021). Recombining Non-Qualitative Reality. *Synthese*, *198*(3), 2273–2295. doi: 10.1007/s11229-019-02204-x

Cowling, S. (2023). Haecceitism. In E. N. Zalta & U. Nodelman (Eds.), *The Stanford encyclopedia of philosophy* (Spring 2023 ed.). Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/spr2023/entries/haecceitism/.

Craver, C. (2017). Stochastic Supervenience. In M. P. Adams, Z. Biener, U. Feest, & J. A. Sullivan (Eds.), *Eppur Si Muove: Doing History and Philosophy of Science with Peter Machamer: A Collection of Essays in Honor of Peter Machamer*. Springer.

Da Costa, N. C. (1980). A Model-Theoretical Approach to Variable Binding Term Operators. In A. Arruda, R. Chuaqui, & N. Da Costa (Eds.), *Mathematical Logic in Latin America* (Vol. 99, p. 133-162). Elsevier. doi: https://doi.org/10.1016/S0049-237X(09)70484-7

Dalla Chiara, M. L., & Toraldo Di Francia, G. (1993). Individuals, Kinds and Names in Physics. In G. Corsi, M. L. Dalla Chiara, & G. C. Ghirardi (Eds.), *Bridging the Gap: Philosophy, Mathematics, and Physics: Lectures on the Foundations of Science* (pp. 261–283). Dordrecht: Springer Netherlands. doi: 10.1007/978-94-011-2496-6_13

Dalla Chiara, M. L., & Toraldo Di Francia, G. (1995). Identity Questions from Quantum Theory. In K. Gavroglu, J. Stachel, & M. W. Wartofsky (Eds.), *Physics, Philosophy, and the Scientific Community: Essays in the philosophy and history of the natural sciences and mathematics. In honor of Robert S. Cohen* (pp. 39–46). Dordrecht: Springer Netherlands. doi: 10.1007/978-94-017-2658-0_3

Daly, C., & Liggins, D. (2010). Do Object-Dependent Properties Threaten Physicalism? *Journal of Philosophy*, *107*(11), 610–614. doi: 10.5840/jphil20101071141

Danto, A. C. (1981). *The Transfiguration of the Commonplace: A Philosophy of Art.* Cambridge, Mass.: Harvard University Press.

Dasgupta, S. (2009). Individuals: An Essay in Revisionary Metaphysics. *Philosophical Studies*, *145*(1), 35–67. doi: 10.1007/s11098-009-9390-x

Dasgupta, S. (2017). Constitutive Explanation. *Philosophical Issues*, *27*(1), 74–97. doi: 10.1111/phis.12102

Davidson, D. (1965). Theories of Meaning and Learnable Languages. In Y. Bar-Hillel (Ed.), *Proceedings of the 1964 International Congress for Logic, Methodology, and Philosophy of Science* (pp. 383–394). Amsterdam: North-Holland.

Davidson, D. (1969). The Individuation of Events. In N. Rescher (Ed.), *Essays in Honor of Carl G. Hempel* (pp. 216–34). Reidel.

De Clercq, R. (2005). A Criterion of Diachronic Identity Based on Locke's Principle. *Metaphysica*, *6*(1), 23–38.

Delgado, L. (2019). Between Singularity and Generality: The Semantic Life of Proper Names. *Linguistics and Philosophy*, *42*(4), 381–417. doi: 10.1007/s10988-018-9251-1

Della Rocca, M. (2005). Two Spheres, Twenty Spheres, and the Identity of Indiscernibles. *Pacific Philosophical Quarterly*, *86*(4), 480–492. doi: 10.1111/j.1468-0114.2005.00238.x

Devitt, M. (2015). Should Proper Names Still Seem So Problematic? In A. Bianchi (Ed.), *On Reference.* Oxford, GB: Oxford University Press UK.

Diekemper, J. (2009). Thisness and Events. *Journal of Philosophy*, *106*(5), 255–276. doi: 10.5840/jphil200910656

Domenech, G., & Holik, F. (2007). A Discussion on Particle Number and Quantum Indistinguishability. *Foundations of Physics*, *37*(6), 855–878. doi: 10.1007/s10701-007-9129-5

Donnellan, K. S. (1966). Reference and Definite Descriptions. *Philosophical Review*, *75*(3), 281–304. doi: 10.2307/2183143

Dorato, M., & Morganti, M. (2013). Grades of Individuality. A Pluralistic View of Identity in Quantum Mechanics and in the Sciences. *Philosophical Studies*, *163*(3), 591–610. doi: 10.1007/s11098-011-9833-z

Dummett, M. (1993). *The Seas of Language*. Oxford University Press.

Eddon, M. (2010). Intrinsicality and Hyperintensionality. *Philosophy and Phenomenological Research*, *82*(2), 314–336. doi: 10.1111/j.1933-1592.2010.00414.x

Effingham, N. (2012). Endurantism and Perdurantism. In R. B. N. Manson (Ed.), *Continuum Companion to Metaphysics* (p. 170).

Esfeld, M. (2004). Quantum Entanglement and a Metaphysics of Relations. *Studies in History and Philosophy of Modern Physics*, *35*(4), 601–617. doi: 10.1016/j.shpsb.2004.04.008

Esfeld, M., & Lam, V. (2008). Moderate Structural Realism About Space-Time. *Synthese*, *160*(1), 27–46. doi: 10.1007/s11229-006-9076-2

Evans, G. (1978). Can There Be Vague Objects? *Analysis*, *38*(4), 208. doi: 10.1093/analys/38.4.208

Fara, D. G. (2015). Names Are Predicates. *Philosophical Review*, *124*(1), 59–117. doi: 10.1215/00318108-2812660

Fine, K. (1977). Properties, Propositions and Sets. *Journal of Philosophical Logic*, *6*(1), 135–191. doi: 10.1007/bf00262054

Fine, K. (1985a). Natural Deduction and Arbitrary Objects. *Journal of Philosophical Logic*, *14*(1), 57–107. doi: 10.1007/bf00542649

Fine, K. (1985b). Plantinga on the Reduction of Possibilist Discourse. In J. E. Tomberlin & P. Van Inwagen (Eds.), *Alvin Plantinga* (pp. 145–186). Dordrecht: Springer Netherlands. doi: 10.1007/978-94-009-5223-2_4

Fine, K. (1985c). *Reasoning with Arbitrary Objects.* New York, NY, USA: Blackwell.

Fine, K. (2000). A Counter-Example to Locke's Thesis. *The Monist*, *83*(3), 357–361. doi: 10.5840/monist200083315

Fine, K. (2012). Guide to Ground. In F. Correia & B. Schnieder (Eds.), *Metaphysical Grounding* (pp. 37–80). Cambridge University Press.

Fine, K., & Tennant, N. (1983). A Defence of Arbitrary Objects. *Aristotelian Society Supplementary Volume*, *57*(1), 55–89. doi: 10.1093/aristoteliansupp/57.1.55

Fodor, J. A. (1987). *Psychosemantics: The Problem of Meaning in the Philosophy of Mind.* MIT Press.

Forbes, G. (1985). *The Metaphysics of Modality.* Clarendon Press.

Fraenkel, A. A., Bar-Hillel, Y., & Lévy, A. (1973). *Foundations of Set Theory.* North-Holland Pub. Co.

Frege, G. (1884/1950). *The Foundations of Arithmetic.* Evanston, Ill.,: Northwestern University Press.

French, S. (1989). Why the Principle of the Identity of Indiscernibles is Not Contingently True Either. *Synthese*, *78*(2), 141–166. doi: 10.1007/bf00869370

French, S. (2019). Identity and Individuality in Quantum Theory. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Winter 2019 ed.). Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/win2019/entries/qt-idind/.

French, S., & Krause, D. (1995). Vague identity and quantum non-individuality. *Analysis*, *55*(1), 20–26. doi: 10.1093/analys/55.1.20

French, S., & Krause, D. (2006). *Identity in Physics: A Historical, Philosophical, and Formal Analysis* (D. Krause, Ed.). New York: Oxford University Press.

French, S., & Ladyman, J. (2003). Remodelling Structural Realism: Quantum Physics and the Metaphysics of Structure. *Synthese*, *136*(1), 31–56. doi: 10.1023/a:1024156116636

French, S., & Redhead, M. (1988). Quantum Physics and the Identity of Indiscernibles. *British Journal for the Philosophy of Science*, *39*(2), 233–246. doi: 10.1093/bjps/39.2.233

Geach, P. T. (1965). *Some Problems about Time*. Oxford University Press.

Geach, P. T. (1967). Identity. *Review of Metaphysics*, *21*(1), 3–12.

Geach, P. T. (2000). *God and the Soul*. London: Routledge and Kegan Paul.

Gödel, K. (1946). Russell's Mathematical Logic. In P. A. Schilpp (Ed.), *The Philosophy of Bertrand Russell, 2nd edition* (pp. 123–154). Evanston, IL: The Library of Living Philosophers, Inc.

Goodman, N. (1965). *Fact, Fiction, and Forecast.* Cambridge, Mass.: Harvard University Press.

Gray, A. (2015). Lexical Individuation and Predicativism About Names. *Thought: A Journal of Philosophy*, *4*(2), 113–123. doi: 10.1002/tht3.164

Haack, R. J. (1978). Davidson on Learnable Languages. *Mind*, *87*(346), 230–249. doi: 10.1093/mind/lxxxvii.2.230

Hacking, I. (1975). The Identity of Indiscernibles. *Journal of Philosophy*, *72*(9), 249–256. doi: 10.2307/2024896

Hawley, K. (2006). Weak Discernibility. *Analysis*, *66*(4), 300–303. doi: 10.1093/analys/66.4.300

Hawley, K. (2009). Identity and Indiscernibility. *Mind*, *118*(469), 101–119. doi: 10.1093/mind/fzn153

Hawthorne, J. (1995). The Bundle Theory of Substance and the Identity of Indiscernibles. *Analysis*, *55*(3), 191–196.

Hawthorne, J. (2003). Identity. In M. J. Loux & D. W. Zimmerman (Eds.), *The Oxford Handbook of Metaphysics* (pp. 99–130). Oxford University Press.

Heil, J. (2003). *From an Ontological Point of View.* Oxford, GB: Oxford University Press.

Hellman, G. (2004). Russell's Absolutism vs. (?) Structuralism. In G. Link (Ed.), *One Hundred Years of Russell's Paradox* (pp. 561–576). Berlin, New York: De Gruyter. doi: doi:10.1515/9783110199680.561

Hellman, G. (2007). Structuralism. In *The Oxford Handbook of Philosophy of Mathematics and Logic.* Oxford University Press. doi: 10.1093/oxfordhb/9780195325928.003.0017

Hempel, C. G., & Oppenheim, P. (1948). Studies in the Logic of Explanation. *Philosophy of Science*, *15*(2), 135–175. doi: 10.1086/286983

Hintikka, J. (1996). *The Principles of Mathematics Revisited.* New York: Cambridge University Press.

Hoffmann-Kolss, V. (2010). *The Metaphysics of Extrinsic Properties.* Ontos-Verlag.

Hoffmann-Kolss, V. (2015). On a Sufficient Condition for Hyperintensionality. *Philosophical Quarterly*, *65*(260), 336–354. doi: 10.1093/pq/pqv021

Hoffmann-Kolss, V. (2019). Defining Qualitative Properties. *Erkenntnis*, *84*(5), 995–1010. doi: 10.1007/s10670-018-9991-x

Hofweber, T. (2005). Supervenience and Object-Dependant Properties. *Journal of Philosophy*, *102*(1), 5–32. doi: 10.5840/jphil2005102128

Hofweber, T. (2015). The Place of Subjects in the Metaphysics of Material Objects. *Dialectica*, *69*(4), 473–490. doi: 10.1111/1746-8361.12121

Horgan, T. (1982). Supervenience and Microphysics. *Pacific Philosophical Quarterly*, *63*(1), 29–43. doi: 10.1111/j.1468-0114.1982.tb00084.x

Horsten, L. (2010). Impredicative Identity Criteria. *Philosophy and Phenomenological Research*, *80*(2), 411–439. doi: 10.1111/j.1933-1592.2010.00325.x

Hoy, R. C. (1984). Inquiry, Intrinsic Properties, and the Identity of Indiscernibles. *Synthese*, *61*(3), 275–297. doi: 10.1007/bf00485055

Jackson, F. (1998). *From Metaphysics to Ethics: A Defence of Conceptual Analysis*. Oxford, GB: Oxford University Press.

Jantzen, B. C. (2011). No Two Entities Without Identity. *Synthese*, *181*(3), 433–450. doi: 10.1007/s11229-010-9717-3

Jantzen, B. C. (2019). Entities Without Identity: A Semantical Dilemma. *Erkenntnis*, *84*(2), 283–308. doi: 10.1007/s10670-017-9958-3

Karofsky, A. (2021). *A Case for Necessitarianism*. New York: Routledge.

Kearns, S., & Magidor, O. (2012). Semantic Sovereignty. *Philosophy and Phenomenological Research*, *85*(2), 322–350. doi: 10.1111/j.1933-1592.2012.00624.x

Keranen, J. (2001). The Identity Problem for Realist Structuralism. *Philosophia Mathematica*, *9*(3), 308–330. doi: 10.1093/philmat/9.3.308

Ketland, J. (2011). Identity and Indiscernibility. *Review of Symbolic Logic*, *4*(2), 171–185. doi: 10.1017/s1755020310000328

Khamara, E. J. (1988). Indiscernibles and the Absolute Theory of Space and Time. *Studia Leibnitiana*, *20*(2), 140–159.

Kim, J. (1982). Psychophysical Supervenience. *Philosophical Studies*, *41*(January), 51–70. doi: 10.1007/bf00353523

King, J. C. (1991). Instantial Terms, Anaphora and Arbitrary Objects. *Philosophical Studies*, *61*(3), 239–265. doi: 10.1007/bf00372355

Kment, B. (2012). Haecceitism, Chance, and Counterfactuals. *Philosophical Review*, *121*(4), 573–609. doi: 10.1215/00318108-1630930

Kment, B. (2014). *Modality and Explanatory Reasoning.* New York: Oxford University Press.

Korbmacher, J., & Schiemer, G. (2018). What Are Structural Properties?? *Philosophia Mathematica*, *26*(3), 295–323. doi: 10.1093/philmat/nkx011

Krause, D. (1992). On a Quasi-Set Theory. *Notre Dame Journal of Formal Logic*, *33*(3), 402–11. doi: 10.1305/ndjfl/1093634404

Krause, D. (2004, January). *The Mathematics of Non-Individuality.* Retrieved from http://philsci-archive.pitt.edu/1592/

Krause, D., & Coelho, A. M. N. (2005). Identity, Indiscernibility, and Philosophical Claims. *Axiomathes*, *15*(2), 191–210. doi: 10.1007/s10516-004-6678-5

Kripke, S. A. (1971). Identity and Necessity. In M. K. Munitz (Ed.), *Identity and Individuation* (pp. 135–164). New York: New York University Press.

Kripke, S. A. (1980). *Naming and Necessity: Lectures Given to the Princeton University Philosophy Colloquium* (D. Byrne & M. Kölbel, Eds.). Cambridge, MA: Harvard University Press.

Ladyman, J. (2005). Mathematical Structuralism and the Identity of Indiscernibles. *Analysis*, *65*(3), 218–221. doi: 10.1111/j.1467-8284.2005.00552.x

Ladyman, J., Linnebo, Ø., & Pettigrew, R. (2012). Identity and Discernibility in Philosophy and Logic. *Review of Symbolic Logic*, *5*(1), 162–186. doi: 10.1017/s1755020311000281

Ladyman, J., & Ross, D. (2007). *Every Thing Must Go: Metaphysics Naturalized* (D. Ross, D. Spurrett, & J. G. Collier, Eds.). New York: Oxford University Press.

Landau, L. D., & Lifschitz. (1959). *Quantum mechanics.* Oxford.

Langton, R., & Lewis, D. K. (1998). Defining 'Intrinsic'. *Philosophy and Phenomenological Research*, *58*(2), 333–345. doi: 10.2307/2653512

Lewis, D. K. (1968). Counterpart Theory and Quantified Modal Logic. *Journal of Philosophy*, *65*(5), 113–126. doi: 10.2307/2024555

Lewis, D. K. (1976). Survival and Identity. In A. O. Rorty (Ed.), *The Identities of Persons* (pp. 17–40). University of California Press.

Lewis, D. K. (1983). Extrinsic Properties. *Philosophical Studies*, *44*(2), 197–200. doi: 10.1007/bf00354100

Lewis, D. K. (1986). *On the Plurality of Worlds.* Malden, Mass.: Wiley-Blackwell.

Lewis, D. K. (1988). Rearrangement of Particles: Reply to Lowe. *Analysis*, *48*(2), 65–72. doi: 10.1093/analys/48.2.65

Locke, D. (2012). Quidditism Without Quiddities. *Philosophical Studies*, *160*(3), 345–363. doi: 10.1007/s11098-011-9722-5

Locke, J. (1689). *An Essay Concerning Human Understanding* (P. Phemister, Ed.). New York: Oxford University Press.

Loux, M. J., & Loux, W. J. (1978). *Substance and Attribute: A Study in Ontology*. Springer Verlag.

Lowe, E. J. (1989). Impredicative Identity Criteria and Davidson's Criterion of Event Identity. *Analysis*, *49*(4), 178–181. doi: 10.1093/analys/49.4.178

Lowe, E. J. (1998). *The Possibility of Metaphysics: Substance, Identity, and Time*. Oxford, GB: Clarendon Press.

Lowe, E. J. (2003). Individuation. In M. J. Loux & D. W. Zimmerman (Eds.), *The Oxford Handbook of Metaphysics.* Oxford University Press.

Lowe, E. J. (2012). The probable simplicity of personal identity. In G. Gasser & M. Stefan (Eds.), *Personal Identity: Complex or Simple?* (p. 137–155). Cambridge University Press. doi: 10.1017/CBO9781139028486.010

Lowe, E. J. (2016). Non-Individuals. In T. Pradeu & A. Guay (Eds.), *Individuals Across the Sciences.* Oxford University Press.

MacBride, F. (2006). What Constitutes the Numerical Diversity of Mathematical Objects? *Analysis*, *66*(1), 63–69. doi: 10.1093/analys/66.1.63

Mackie, P. (2006). *How Things Might Have Been: Individuals, Kinds, and Essential Properties.* Oxford University Press.

Magidor, O. (2011). Arguments by Leibniz's Law in Metaphysics. *Philosophy Compass*, *6*(3), 180–195. doi: 10.1111/j.1747-9991.2010.00379.x

Markosian, N. (1998). Simples. *Australasian Journal of Philosophy*, *76*(2), 213–228. doi: 10.1080/00048409812348361

Marshall, D., & Weatherson, B. (2018). Intrinsic vs. Extrinsic Properties. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Spring 2018 ed.). Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/spr2018/entries/intrinsic-extrinsic/.

Martino, E. (2001). Arbitrary Reference in Mathematical Reasoning. *Topoi*, *20*(1), 65–77. doi: 10.1023/a:1010613027576

Melia, J. (1999). Holes, Haecceitism and Two Conceptions of Determinism. *British Journal for the Philosophy of Science*, *50*(4), 639–64. doi: 10.1093/bjps/50.4.639

Menzel, C. (2022). Actualism. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Summer 2022 ed.). Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/sum2022/entries/actualism/.

Merricks, T. (1997). Fission and Personal Identity Over Time. *Philosophical Studies*, *88*(2), 163–186. doi: 10.1023/a:1004210420052

Montero, B. (2013). Must Physicalism Imply the Supervenience of the Mental on the Physical? *Journal of Philosophy*, *110*(2), 93–110. doi: 10.5840/jphil2013110240

Moreland, J. P. (2001). *Universals.* Routledge.

Moreland, J. P. (2013). Exemplification and Constituent Realism: A Clarification and Modest Defense. *Axiomathes*, *23*(2), 247–259. doi: 10.1007/s10516-011-9148-x

Muller, F. A., & Seevinck, M. P. (2009). Discerning Elementary Particles. *Philosophy of Science*, *76*(2), 179–200. doi: 10.1086/647486

Nagel, G. (1976). The Identity of Indiscernibles. *Journal of Philosophy*, *73*(2), 45–50. doi: 10.2307/2025460

Noonan, H. W. (1980). *Objects and Identity: An Examination of the Relative Identity Thesis and its Consequences.* The Hague: Martinus Nijhoff.

Orilia, F. (2006). Identity Across Time and Stories. In A. Bottani & R. Davies (Eds.), *Modes of Existence: Papers in Ontology and Philosophical Logic* (pp. 191–220). Ontos Verlag.

Owen, D. (1991). Locke on Real Essence. *History of Philosophy Quarterly*, *8*(2), 105–118.

Parkinson, G. H. R. (1965). *Logic and Reality in Leibniz's Metaphysics.* New York: Garland.

Parkinson, G. H. R. (1981). Kant as a Critic of Leibniz. The Amphiboly of Concepts of Reflection. *Revue Internationale de Philosophie*, *35*(136/137), 302.

Plantinga, A. (1974). *The Nature of Necessity.* Oxford, England: Clarendon Press.

Plantinga, A. (2003). Transworld Identity or Worldbound Individuals? In *Essays in the Metaphysics of Modality.* Oxford University Press. doi: 10.1093/0195103769.003.0004

Pooley, O. (2005). Points, Particles, and Structural Realism. In D. Rickles, S. French, & J. T. Saatsi (Eds.), *The Structural Foundations of Quantum Gravity* (pp. 83–120). Oxford University Press.

Post, E. (1963). Individuality and Physics. *The Listener*, *10 October*(3), 534–537.

Priest, G. (1997). Inconsistent Models of Arithmetic Part I: Finite Models. *Journal of Philosophical Logic*, *26*(2), 223–235. doi: 10.1023/a:1004251506208

Priest, G. (2008). Contingent Identity in Modal Logic. In *An Introduction to Non-Classical Logic: From If to Is* (2nd ed., pp. 367–383). Cambridge University Press. doi: 10.1017/CBO9780511801174.022

Putnam, H. (1975). The Meaning of 'Meaning'. *Minnesota Studies in the Philosophy of Science*, *7*, 131–193.

Quine, W. V. O. (1976). Grades of Discriminability. *Journal of Philosophy*, *73*, 113–6.

Quine, W. V. O. (1985). Events and Reification. In E. Lepore & B. McLaughlin (Eds.), *Actions and Events: Perspectives on the Philosophy of Davidson* (pp. 162–71). Blackwell.

Quinton, A. (1973). *The Nature of Things.* Boston,: Routledge and Kegan Paul.

Read, S. (1997). Completeness and Categoricity: Frege, Gödel and Model Theory. *History and Philosophy of Logic*, *18*(2), 79–93. doi: 10.1080/01445349708837277

Rescher, N. (1967). *The Philosophy of Leibniz.* Prentice-Hall.

Rieger, A. (2000). An Argument for Finsler-Aczel Set Theory. *Mind*, *109*(434), 241–253.

Rinard, S. (2021). Pragmatic Skepticism. *Philosophy and Phenomenological Research*, *104*(2), 434–453. doi: 10.1111/phpr.12779

Robinson, A. (1961). Non-standard analysis. *Proceedings of the Royal Academy of Sciences, Amsterdam*, *64*(ser. A), 432–440.

Robinson, A. (1966). *Non-standard analysis.* Amsterdam, The Netherlands: North-Holland.

Robinson, D. (1989). Matter, Motion, and Humean Supervenience. *Australasian Journal of Philosophy*, *67*(4), 394–409. doi: 10.1080/00048408912343921

Rodriguez-Pereyra, G. (2002). *Resemblance Nominalism: A Solution to the Problem of Universals.* New York: Clarendon Press.

Rodriguez-Pereyra, G. (2006). How Not to Trivialise the Identity of Indiscernibles. In P. F. Strawson & A. Chakrabarti (Eds.), *Concepts, Properties and Qualities.* Ashgate.

Rodriguez-Pereyra, G. (2017). Indiscernible Universals. *Inquiry: An Interdisciplinary Journal of Philosophy*, *60*(6), 604–624. doi: 10.1080/0020174x.2016.1210541

Rodriguez-Pereyra, G. (2018). The Principles of Contradiction, Sufficient Reason, and Identity of Indiscernibles. In M. R. Antognazza (Ed.), *Oxford Handbook of Leibniz.* Oxford University Press.

Rosen, G. (2010). Metaphysical Dependence: Grounding and Reduction. In B. Hale & A. Hoffmann (Eds.), *Modality: Metaphysics, Logic, and Epistemology* (pp. 109–135). Oxford University Press.

Rosenkrantz, G. S. (1979). The Pure and the Impure. *Logique Et Analyse*, *22*(88), 515.

Roy, T. (1995). In Defense of Linguistic Ersatzism. *Philosophical Studies*, *80*(3), 217–242. doi: 10.1007/bf00990584

Ruffing, M., Almeida, G. A. D., Terra, R. R., & Rohden, V. (2008). Kant's Argument in the Amphiboly. In M. Ruffing, G. A. D. Almeida, R. R. Terra, & V. Rohden (Eds.), *Recht Und Frieden in der Philosophie Kants. Law and Peace in Kant's Philosophy: Akten des X. Internationalen Kant-Kongresses.* Walter de Gruyter.

Russell, B. (1992). *The Philosophy of Leibniz.* Routledge.

Saunders, S. (2003). Physics and Leibniz's Principles. In K. Brading & E. Castellani (Eds.), *Symmetries in Physics: Philosophical Reflections* (pp. 289–307). Cambridge University Press.

Saunders, S. (2006). Are Quantum Particles Objects? *Analysis*, *66*(1), 52–63. doi: 10.1093/analys/66.1.52

Saunders, S., & Muller, F. A. (2008). Discerning fermions. *British Journal for the Philosophy of Science*, *59*(3), 499–548. doi: 10.1093/bjps/axn027

Schaffer, J. (2009). On What Grounds What. In D. Manley, D. J. Chalmers, & R. Wasserman (Eds.), *Metametaphysics: New Essays on the Foundations of Ontology* (pp. 347–383). Oxford University Press.

Schaffer, J. (2010). The Least Discerning and Most Promiscuous Truthmaker. *Philosophical Quarterly*, *60*(239), 307–324. doi: 10.1111/phiq.2010.60.issue-239

Schrödinger, E. (1953). What is matter? *Scientific American*, *September*, 52–57.

Schrödinger, E. (1963). *My View of the World*. Cambridge: University Press.

Shapiro, S. (1997). *Philosophy of Mathematics: Structure and Ontology*. New York: Oxford University Press.

Sider, T. (2002). The Ersatz Pluriverse. *Journal of Philosophy*, *99*(6), 279–315.

Simmons, B. (2020). Impure Concepts and Non-Qualitative Properties. *Synthese*, *197*(7), 3065–3086. doi: 10.1007/s11229-018-1871-z

Skow, B. (2008). Haecceitism, Anti-Haecceitism, and Possible Worlds: A Case Study. *Philosophical Quarterly*, *58*(230), 97–107. doi: 10.1111/j.1467-9213.2007.516.x

Smid, J. (2021). Mind the Gap: The Space Between Coincidence and Colocation. *Journal of Philosophical Logic*, *51*(1), 57–73. doi: 10.1007/s10992-021-09609-5

Stachel, J. (2005). Structure, Individuality, and Quantum Gravity. In D. Rickles, S. French, & J. T. Saatsi (Eds.), *The Structural Foundations of Quantum Gravity.* Clarendon Press.

Stalnaker, R. C. (1984). *Inquiry.* Cambridge University Press.

Strawson, P. (1959). *Individuals: An Essay in Descriptive Metaphysics* (W. Wang, Ed.). London, England: Routledge.

Stroud, B. (1984). *The Significance of Philosophical Scepticism.* New York: Oxford University Press.

Swinburne, R. (1995). Thisness. *Australasian Journal of Philosophy*, *73*(3), 389–400. doi: 10.1080/00048409512346721

Tarski, A., & Woodger, J. H. (Eds.). ([1936] 1983). *Logic, Semantics, Metamathematics: Papers From 1923 to 1938.* New York, NY, USA: Hackett Publishing Company.

Turner, J. (2016). Can We Do Without Fundamental Individuals? No. In E. Barnes (Ed.), *Current Controversies in Metaphysics (1st ed.).* Routledge.

Van Bendegem, J. P. (1999). Why the Largest Number Imaginable is Still a Finite Number. *Logique Et Analyse*, *42*(165-166).

Van Inwagen, P. (1981). The Doctrine of Arbitrary Undetached Parts. *Pacific Philosophical Quarterly*, *62*(2), 123–137. doi: 10.1111/j.1468-0114.1981.tb00051.x

Van Inwagen, P. (1990). Four-Dimensional Objects. *Noûs*, *24*(2), 245–255.

Varzi, A. C. (2003). Perdurantism, Universalism and Quantifiers. *Australasian Journal of Philosophy*, *81*(2), 208–215. doi: 10.1080/713659632

Wenmackers, S. (2016). Infinitesimal Probabilities. In R. Pettigrew & J. Weisberg (Eds.), *The Open Handbook of Formal Epistemology* (pp. 199–265). PhilPapers Foundation.

Weyl, H. (1950). *The Theory of Groups and Quantum Mechanics.* Dover Publications.

Wiggins, D. (1965). Identity Statements. In R. J. Butler (Ed.), *Analytic Philosophy, 2nd edition* (pp. 40–71). Blackwell.

Wiggins, D. (2001). *Sameness and Substance Renewed* (D. Wiggins, Ed.). New York: Cambridge University Press.

Williams, D. C. (1966). *Principles of Empirical Realism: Philosophical Essays.* Springfield, Ill.,: C.C. Thomas.

Williamson, T. (1990). *Identity and Discrimination.* Cambridge, Mass., USA: Wiley-Blackwell.

Williamson, T. (1998). Bare Possibilia. *Erkenntnis*, *48*(2-3), 257–73. doi: 10.1023/a:1005331819843

Williamson, T. (2013). *Modal Logic as Metaphysics.* Oxford, England: Oxford University Press.

Williamson, T. (2022). The Linguistic Turn and the Conceptual Turn. In *The Philosophy of Philosophy* (pp. 12–24). Wiley.

Woods, J. (2014). Logical Indefinites. *Logique Et Analyse – Special Issue Edited by Julien Murzi and Massimiliano Carrara*, *227*, 277–307.

Wright, C. (2004). Warrant for Nothing (and Foundations for Free)? *Aristotelian Society Supplementary Volume*, *78*(1), 167–212. doi: 10.1111/j.0309-7013.2004.00121.x

Wüthrich, C. (2009). Challenging the Spacetime Structuralist. *Philosophy of Science*, *76*(5), 1039–1051. doi: 10.1086/605825

Yablo, S. (1993). Is Conceivability a Guide to Possibility? *Philosophy and Phenomenological Research*, *53*(1), 1–42. doi: 10.2307/2108052