Synthese Library 374

Jiri Benovsky

Meta-metaphysics

On Metaphysical Equivalence, Primitiveness, and Theory Choice

Springer

Synthese Library

Studies in Epistemology, Logic, Methodology, and Philosophy of Science

Volume 374

More information about this series at http://www.springer.com/series/6607

Jiri Benovsky

Meta-metaphysics

On Metaphysical Equivalence, Primitiveness, and Theory Choice

Springer

Jiri Benovsky
Fribourg
Switzerland

Synthese Library
ISBN 978-3-319-25332-9 ISBN 978-3-319-25334-3 (eBook)
DOI 10.1007/978-3-319-25334-3

Library of Congress Control Number: 2016933112

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG Switzerland

Preface

Metaphysical theories are beautiful. I mean it literally. At the end of this book, I will defend the view that metaphysical theories possess aesthetic properties and that these play a crucial role when it comes to theory evaluation and theory choice.

But this is the end of a long journey—a journey that is perhaps more important than the destination. Before we get there, the philosophical path I propose to follow starts with three discussions of metaphysical equivalence. I shall begin with Relationism and Substantivalism about time, and with the Bundle theory and Substratum theory of material objects, and in both cases I will argue for a kind of equivalence between these traditional enemies—thus, we shall have two original examples of metaphysical equivalence.

Second, I will discuss the case of Perdurantism and Endurantism, which some have thought of as a good example of metaphysical equivalence as well. As we shall see, however, this is correct only to some extent, and there does remain room for substantive—as opposed to purely verbal—disagreement. As my examination of this debate will show, there is room for the metaontological/methodological view that a debate is largely, but not entirely, merely verbal and that some parts of it are substantive, and decidable by philosophical methods. Thus, we shall have an original example of *partial* metaphysical equivalence.

I will then focus on the Presentism versus Eternalism controversy, and argue against the claims of equivalence between these two views that have been raised in the literature. I believe that these views are not equivalent in any interesting sense, and more importantly, I believe that those who make claims of equivalence concerning them do so because they use a flawed methodology when doing metaontology. Thus, we shall have an example (as well as others) of metaphysical non-equivalence.

So, here is the first thing that I wish to achieve in this book, and that will stem from the discussions above: articulate a metaontological view which emphasizes that when asking the question "Are metaphysical debates substantive or verbal?" the correct answer is "It depends." Some debates are substantive, some debates are merely verbal, sometimes it is true that a problem or a question can be formulated in

equally good frameworks where there is no fact of the matter as to which one is correct or where we just cannot know it. Furthermore, it is possible, and it is true in the case of the persistence debate, that inside a debate some points are merely verbal while other are places of substantive disagreement. A general methodological point will arise: the best way to do meta-metaphysics is to do first-level metaphysics.

A second thing I will elaborate at this point concerns primitiveness. Indeed, my way of arguing for or against any equivalence claims largely depends on the nature of primitives and on the role they play in each of the theories involved. In general, in most metaphysical debates a lot depends on primitives—indeed, metaphysical theories heavily rely on the use of primitives that they typically appeal to. So, I will emphasize here the utmost importance of primitives in the construction of metaphysical theories and in the subsequent evaluation of them. I will claim that almost all of the explanatory power of metaphysical theories comes from their primitives, and I will scrutinize the notion of "power" and "explanatory". All together, these points will naturally lead me to defend a global view on the nature of the metaphysical enterprise: what is at stake in metaphysics is to find out not just what there is or what there is not, but what is more fundamental than what—to find out what are the best primitives.

Armed with this understanding of the way metaphysical theories work, I shall then raise the simple but complicated question: how to make a choice between competing metaphysical theories? If two theories are equivalent, then perhaps we do not need to make a choice. But what about all the other cases of non-equivalent "equally good" theories? I shall use some of the theories discussed above as examples and I shall examine some traditional meta-theoretical criteria for theory choice (various kinds of simplicity, compatibility with physics, compatibility with intuitions, explanatory power, internal consistency,…) only to show that they do not allow us to make a choice—that is following one or more of these criteria will not help us to tell which theory is preferable to the others. We will see that even in cases of non-equivalent theories, metaphysicians can find themselves in a situation where it is far from clear how to make a choice between competing theories, or even that such a choice can be made.

Among the meta-theoretical criteria for theory evaluation and theory choice, compatibility with our intuitions strikes me as being especially important in metaphysics. Metaphysical theories are often counter-intuitive. But they are often also strongly supported and motivated by intuitions. One way or another, the link between intuitions and metaphysics is a strong and important one, and there is hardly any metaphysical discussion where intuitions do not play a crucial role. I will focus on a particular kind of such intuitions, namely those that come, at least partly, from experience. There seems to be a route from experience to metaphysics, and I shall examine it carefully. At the end of the day, I shall argue that this route is a treacherous one, and phenomenological considerations are in fact orthogonal to the allegedly 'corresponding' metaphysical claims.

At this point of this philosophical journey, one may want to sit back and reflect on what the above considerations—if correct—imply. If neither intuitions nor other meta-theoretical criteria can help us in deciding between competing non-equivalent

metaphysical theories, how then shall we make that choice? Here again, considerations about the theories' explanatory power, coming mostly from their primitives, will play a crucial role in the sense that it will appear that there are important cases of non-equivalent, competing, but equally good, theories (my main example will be the Tropes vs. Universals vs. Nominalism debate). What to do then?

This is where I shall argue that metaphysical theories possess aesthetic properties—grounded in non-aesthetic properties—and that these play a crucial role in theory choice and evaluation. Indeed, it seems that the aesthetic properties of a theory can be appealed to when it comes to preferring one theory over another. In short, the view at hand is that metaphysical theories are beautiful and that contemplating their beauty is what drives us to prefer one to another.

This view, as well as all the meta-metaphysical considerations discussed throughout the book, will naturally lead me to a form of anti-realism, and at the end of the journey I shall sit down to rest, and to offer reasons to think better of the kind of anti-realist view I propose to embrace.

Acknowledgments

I owe many thanks to many temporal parts of many people.

First of all, I would like to thank Laurie Paul who helped me a lot to clarify my own ideas and explore directions that proved to be fruitful; her help was invaluable. For very helpful comments, feedback, discussions, and email conversations, I would like to thank Mike Bertrand, Eric Cavallero, Alix Cohen, Fabrice Correia, Damiano Costa, Barry Dainton, Maksymilian DelMar, Coralie Dorsaz, Fabian Dorsch, Michael Esfeld, Akiko Frischut, Robert Fudge, Amanda Garcia, Richard Glauser, Ghislain Guigon, Mark Heller, Rob Hopkins, Thomas Jacobi, Max Kistler, Kathrin Koslicki, Baptiste Le Bihan, Robin Le Poidevin, Vincent Lam, Stephan Leuenberger, Dan López de Sa, Jonathan Lowe, Olivier Massin, Trenton Merricks, Anne Meylan, Kevin Mulligan, Nathan Oaklander, Alain Pé-Curto, Roger Pouivet, Gonzalo Rodriguez-Pereyra, Sven Rosenkranz, Martine Nida-Ruemelin, Ted Sider, Jonathan Schaffer, Luc Schneider, Daniel Schulthess, Gianfranco Soldati, Cain Todd, Achille Varzi and John Zeimbekis.

Special thanks go to Lynda, not only for her philosophical feedback on parts of this book, but also for her love and moral support during the time I was writing it.

This book brings together ideas I have been working on in the 6 six years. Some of these ideas have already appeared as journal articles, and I based some of the chapters of this book on them. They are: 'Tropes or universals: how (not) to make one's choice', 2014, *Metaphilosophy*; 'Primitiveness, metaontology, and explanatory power', 2013, *Dialogue: Canadian Philosophical Review*; 'From experience to metaphysics: on experience-based intuitions and their role in metaphysics', Noûs, 2015, 49:4, p. 684–697; 'Philosophical theories, aesthetic value, and theory choice', 2013, *The Journal of Value Inquiry*; 'Aesthetic Supervenience vs. Aesthetic Grounding', 2012, *Estetika: The Central European Journal of Aesthetics*; 'The relationist and substantivalist theories of time: foes or friends?', 2011, *European Journal of Philosophy*; 'Endurance, perdurance, and metaontology', 2011, *Northern European Journal of Philosophy (Sats)*; 'Relational and substantival

ontologies, and the nature and the role of primitives in ontological theories', 2010, *Erkenntnis*; 'The bundle theory and the substratum theory: deadly enemies or twin brothers?', 2008, *Philosophical Studies*; 'On Presentist Perdurantism', 2007, *Northern European Journal of Philosophy (Sats)*.

Contents

Part I
On Equivalence and Primitiveness

Chapter 1
Equivalent Metaphysical Theories

Abstract In this chapter, I discuss various cases of metaphysical equivalence. I focus on the case of the Bundle theory and Substratum theory of material objects and the case of Relationism and Substantivalism about time, and in both cases I argue for a kind of equivalence between these traditional enemies. I explore several versions of the Bundle theory and the Substratum theory. First, I examine different versions of the Bundle theory with tropes and compare them to the Substratum theory with tropes by going through various standard objections and arguing for a tu quoque in all cases. Emphasizing the theoretical role of the substratum and of the relation of compresence, I defend the claim that these views are equivalent for all theoretical purposes. I then examine two different versions of the Bundle theory with universals, and show that one of them is, here again, equivalent to the Substratum theory with universals, by examining how both views face the famous objection from Identity of Indiscernibles in a completely parallel way. I then examine other versions of these views that are not equivalent to the others. I then turn my attention to Relationism and Substantivalism about time. In this chapter, I try to make some progress with respect to the debate between these two views, and I do this mainly by examining the strategies they use to face the possibilities of 'empty time' and 'time without change'. As we shall see, the two allegedly very different rival views are much less different than has been thought: their structure is extremely similar, their strategies are extremely similar, and they can both face the possibilities of 'empty time' and 'time without change' in the same way. Thus, I argue in favour of a certain kind of equivalence between these two views as well.

§1. There are different ways one can understand what metaphysical equivalence is, and depending on the way one chooses to go, different particular claims of metaphysical equivalence will arise. In this chapter, as well as in the next two chapters, I will *not* first explain which way I have chosen to go, and then apply the methodology to various particular cases. I shall rather proceed the other way around: starting from first-level, non-meta-metaphysical considerations, I will argue for claims of metaphysical equivalence in the case of several particular metaphysical theories (in this chapter), for claims of partial metaphysical equivalence in other cases (in Chap. 2), and for claims of metaphysical non-equivalence in yet other cases (in Chap. 3). Even

J. Benovsky, *Meta-metaphysics*, Synthese Library 374,
DOI 10.1007/978-3-319-25334-3_1

though I will, of course, (have to) disseminate some of my methodology already on the go, I will explicitly explain it only after all this work is done along with some other meta-metaphysical and methodological considerations (partly in Chap. 3, and mostly in Chap. 4). In short, this is "learning by doing": we shall learn about the correct methodology by seeing how it works in various types of cases.

But, since I don't want to frustrate you from the start by saying only what I will *not* do without giving the details of my methodology, I can provide a slogan: *metaphysical theories are equivalent if they do the same job in the same way* (please, do bear in mind that this is a slogan and not a definition). Thus, I will be interested in how metaphysical theories *work*—that is, in their *function* or *explanatory power*, and in where this power comes from. I will claim that some metaphysical theories are equivalent because they do their job in the same way, using the same—often primitive—tools. This is different from some other claims of metaphysical equivalence which focus on 'translatability' of one view into another and on their respective expressive powers (while, of course, there are important connections and some similarities). Various types of such claims can be found in Callender (2000), Dorato (2006), Lombard (1999), McCall and Lowe (2003), Miller (2005a, b), and perhaps most influentially in Hirsch (2005, 2007, 2008). I shall critically discuss these views in the relevant places in the next three chapters, and compare them to my own, mostly in order to contrast them with my view and help me to articulate it.

The question of metaphysical equivalence has a strong relation with a general debate in meta-metaphysics, which can be seen as a choice between two possible answers that have been given to the question of whether metaphysical debates are really substantive or whether they are merely verbal. This question has received various answers, giving rise to some 'extreme' and some 'moderate' views. On the two opposite sides of the spectrum of the debate lie two 'extreme' views: first, the *realist view* recently advocated for by Sider (2001b, 2007, 2008, 2011) defends the claim that metaphysical disputes *are* substantive and that metaphysical questions have objective answers, while the *sceptical anti-realist view* defended in the recent debate in different ways by Chalmers (2008) and Yablo (2008) claims that metaphysical questions do not have objective answers, that they can be formulated and answered in different frameworks, and that there is no fact of the matter as to which framework is correct—thus, metaphysical claims lack truth-value. In between these two 'extreme' views lie two 'moderate' ones. Bennett (2008) defends an *epistemicist view* that claims that at least some metaphysical questions have genuine objective answers but that often we cannot discover them and that consequently there is often little reason or no reason at all to go for one side rather than the other, and Hirsch (2005, 2007, 2008) defends a *moderate anti-realist view* that claims that many metaphysical debates are merely verbal disputes where the disputants seem to claim different things but in fact they are making the same claims only formulated in different ways, or different languages. In many ways, while my own view is different, it has sympathies with these two latter 'moderate' views. Here is another slogan I can put forward to express where I think moderation lies: *when asking whether metaphysical disputes are merely verbal or whether they are genuine, the answer is "It depends"*. It depends on how the metaphysical theories at hand do their job.

That's enough slogans, let us now start the real work.

§2. The first case I will be interested in is the debate between the bundle theory and the substratum theory. In this chapter, I will explore several versions of these two views and compare them with the rather surprising result that it seems to be true that they *are* equivalent. In order to see whether this is correct or not, I will go through several steps: first, I examine different versions of the bundle theory with tropes and compare them to the substratum theory with tropes by going through various standard objections and arguing for a *tu quoque* in all cases. Emphasizing the theoretical role of the substratum and of the relation of compresence, I defend the claim that these views are equivalent for all theoretical purposes. I then examine two different versions of the bundle theory with universals, and show that one of them is, here again, equivalent to the substratum theory with universals, by examining how both views face the famous objection from Identity of Indiscernibles in a completely parallel way. It is only the second, quite extreme and puzzling, version of the bundle theory with universals that is not equivalent to any other view; and the diagnosis of why this is so will show just how unpalatable the view is. Similarly, only a not-so-palatable version of the substratum theory is genuinely different from the other views; and here again it's precisely what makes it different that makes it less appealing.

The substratum theory (also called 'the substance-attribute view') can be put as a claim about what the relationship between a particular and its properties is: there are particulars and there are the properties that are exemplified, instantiated, or had by the particulars who are conceived of as being the bearers of those properties. Such a bearer of properties, which has its identity independently of the properties which it bears, is often called "an underlying subject", a "substratum", or a "bare particular". An object like a table or a person is thus made out of two different kinds of components: properties and a substratum that supports them and glues them together in order to make up an object.

In contrast, the bundle theory denies the existence and the need for a substratum: as fundamental components of reality, there are only properties. Take my neighbour Cyrano: he is of a certain age, he has a big nose, he has such and such a height, and so on. And this is all there is to know, and all there is to *be* Cyrano—his properties. On this view, an object is then taken to be a bundle (a cluster, a bunch, …) of its properties. There is no need and no room in the bundle theory for two kinds of components to make up objects, rather, they are just bundles of properties which are the ultimate constituents of reality, and which are held together (glued together in order to make up an object) by a special property (an n-adic relation, where n is the number of properties of the object) often called "compresence" (following Russell, who meant by this label something like "simultaneous presence"; however, when speaking here about the bundle theory and while keeping the term "compresence" it is not only and not specifically Russell's view that I will have in mind).

Depending on how one conceives of the compresence relation (the 'bundling' relation) and of the nature of properties, the bundle theory comes in different versions. First, as combined with trope theory:

1. the bundle theory with properties as tropes, where the compresence relation is one and the same numerically identical relation for all objects
2. the bundle theory with properties as tropes, where the compresence relation is one and the same numerically identical, variably polyadic relation for all objects
3. the bundle theory with properties as tropes, where there are distinct compresence relations (one per object)

and second, as combined with the view that properties are universals:

4. the bundle theory with properties as universals, where the compresence relation is one and the same numerically identical relation for all objects
5. the bundle theory with properties as universals, where there are distinct compresence relations (one per object)
6. the bundle theory with properties as universals, where the compresence relation is one and the same numerically identical, variably polyadic relation for all objects

I will discuss these options in turn, in the order in which they appear above (the really interesting ones being 3, 5, and 6). What about the substratum theory? Is there also such a table to be drawn? Not really: of course, the substratum theory comes in different versions depending on whether properties are conceived of as tropes or as universals, but there are not different possibilities for the substratum as there are for the compresence relation—the substratum is, by definition, numerically different in different particular objects.

§3. (1) and (2) are not really available options. First, simply because tropes, unlike universals, cannot multiply occur in different objects and so it is not possible here that one and the same compresence relation does its bundling work in different objects. Second, (1) is unavailable for another obvious reason: not all objects have the same number of properties, and so something like a variably polyadic relation (like "x, y, z, … are compresent with one another") is required. (This will also be true for (4).)

With respect to the compresence relation, the only really interesting position for the bundle theory with tropes is (3). Let me now compare this view, BTT (Bundle Theory with Tropes) to its alleged opponent STT (Substratum Theory with Tropes).

To make up objects out of properties, BTT uses a bundling relation that goes around under different names like "compresence" (the term that I will be using), "consubstatiation", "co-instatiation", "togetherness", "collocation", etc. The abundance of labels does *not* reflect an abundance of different analysis of what this relation is; rather, the compresence relation is usually taken as unanalyzable and ontologically primitive. It is thus defined and individuated not by its nature or intrinsic features, of which we are not told much by BTT, but rather by its theoretical role: it is a *unifying device,*[1] a device that takes properties to make up objects.

[1]One could also say 'tying device' or 'object-making device'. L.A. Paul (forthcoming_b) puts forward a "mereological bundle theory" where "properties are literally objects and parts of objects, and properties are bundled using the composition relation" (§2). This version of a bundle theory thus does not use a primitive bundling relation such as "compresence" or similar, but it appeals to a primitive relation of mereological composition (i.e. a primitive notion of "is a proper part of").

Compare this to STT (and see how easy it is for me to write an almost exact paraphrase of the preceding paragraph here). To make up objects, STT uses properties and a bearer of properties that goes around under different names like "substratum" (the term that I will be using), "naked particular", "bare particular", "thin particular", "substance" (but be careful about this one), etc. The abundance of labels does *not* reflect an abundance of different analysis of what this bearer is; rather, the substratum is usually taken as unanalyzable and ontologically primitive. It is thus defined and individuated not by its nature or intrinsic features, of which we are not told much by STT, but rather by its theoretical role: it is a *unifying device*, a device that takes properties to make objects.

Both BTT and STT thus have a unifying device, a primitive and under-defined one, an entity whose purpose is to tie or glue together properties of a single object. Paraphrasing Locke, in *both* cases this unifying device is a "we-know-not-what" … but it is a "we-know-what-it-does", that is, we know its theoretical role.[2]

One often asks, as an objection to STT: "In virtue of what is a substratum distinct from another substratum? No attributes or properties can distinguish between them!" But the very same question can be asked about compresence: "*Tu quoque*: In virtue of what is one compresence relation (involved in the bundling of an object A) distinct from another (involved in the bundling of another object B)? No attributes or properties can distinguish between them." Both views answer these questions by a primitivist claim.

It is by reflecting on the preceding that it struck me for the first time that I actually don't really see the difference between BTT and STT. But of course much more needs to be done in order to even start to justify any kind of equivalence claim between the two views. I shall do this by first examining a possible difference in the status of the unifying device in BTT and STT (and see that there isn't any), and then by examining some objections to STT and BTT where the unifying device plays a crucial role, and I will argue for a *tu quoque* in all cases.

§4. A possible difference between BTT and STT could be that compresence is just one among other elements of the bundle (just one among the properties of an object), while a substratum has to be considered apart from the properties it bears. So compresence and the substratum don't have the same status, they do *not* play the same theoretical role in the composition of an object.

But this is not true, as Ehring (2001) shows: The properties included in a bundle are compresent. The compresence relation, however, is not a member of the bundle like the other properties and relations because if we included compresence in the bundle,

[2]Compare to what Peter Simons puts as an objection to BTT (my italics): "One possibility is that compresence is neither a binary (gluing two tropes) nor a ternary (gluing two tropes and a place) relation but one with many more terms, as many as there are tropes in the bundle. We may not know what arity this relation has - it might even be infinite - and there might be different arities for different types of concrete independent particular, but there will be such a relation nevertheless. A big drawback this has is that it is hard to see what explanatory force this has. *All we are saying is that a bundle of tropes is held together by whatever relation holds it together*. This is really giving up" (Simons 1994, p. 371).

then it would itself have to be compresent with the other properties: compresence compresent with F, G, H, … But that either makes no sense or leads to an infinite regress. And what it shows is that compresence, exactly as the substratum, has to be considered apart from the other properties of the object; its status as a unifying device is thus different from the other properties and is the same in BTT and STT.

§5. Let us now consider a first of two classic objections to BTT, and examine how STT allegedly avoids them by appealing to a substratum.

This first objection can be found, for instance, in Van Cleve (1985, p. 122): "If a thing were a set of properties, it would be incapable of change. For a thing could change its properties only if the set identical with it could change its members, but that is impossible; no set can change its members." Taking an example of an individual that is supposed to change one of its properties over time, he adds: "[…] what we have is replacement of one individual by another, not change in the properties of one and the same individual" (Van Cleve 1985, p. 124).

The idea here is simple, and quite compelling: if an individual is identified with a bundle of properties, then if one of the properties changes, the bundle is not the same, and so, the individual who is the bundle is not the same—it simply ceased to exist, while another individual has taken its place. So, according to BTT, nothing can undergo change in properties.

How does STT avoid this objection? It doesn't! There is no more genuine change in STT than in BTT: the substratum (the thin particular) does obviously *not* change (since it is propertyless, it cannot undergo qualitative change), and the thick particular (the substratum + the properties it bears) does *not* change either *for exactly the same reason* as in the case of BTT. There is no change, but there is replacement of one thick particular by another: take a thick particular at t_1 and at a later time t_2 when it went through some qualitative change, it simply is something different.

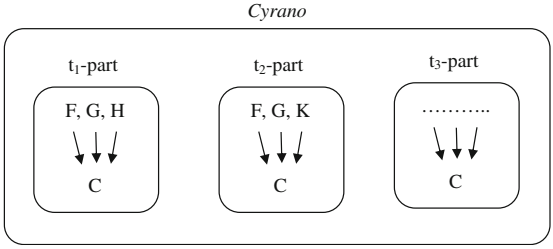
Of course, the STT theorist will object: "I have something that you, BTT, don't have: a substratum that remains the same over different times, and this guarantees me that the individual, while changing its properties, is the same individual."

But if this were an acceptable reply here, then BTT has at hand exactly the same: "I have a compresence relation that remains the same over different times, and this guarantees me that the individual, while changing its properties, is the same individual. You have your primitive unifying device that does the job, I have mine."

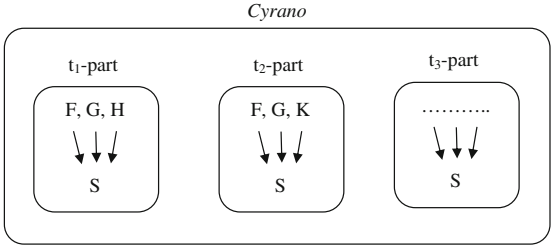
So it seems that both views can handle the objection in the same way by appealing to the unifying device, and that calling the device different names ("compresence" or "substratum") does not change anything since both unifying devices just play the same role in the same way. To be more precise about this, let us examine two main strategies for facing the problem of persistence through time, namely a version of perdurantism and a version of endurantism, and see whether there is any difference between the use of a substratum or the compresence relation. (There isn't.)

§6. A traditional version of perdurantism (the 4D worm view) for BTT is a *bundle-bundle* theory. Think again of my neighbour Cyrano, as a case of an object changing in intrinsic properties over time: Cyrano has, at t_1, a big nose but undergoes plastic surgery, and has at a later time t_2 a small nose. According to perdurantism, Cyrano is an aggregate of his temporal parts, which are numerically and qualitatively

distinct objects. The perdurantist account of change thus claims that Cyrano changes over time by having different temporal parts at different times. In terms of BTT, this picture looks as follows (where "C" stands for "compresence"):



Cyrano is here simply a bundle of momentary bundles of properties (a bundle of bundles). Now, what about a perdurantist version of STT? It does not look any different from BTT-like perdurantism, except that I had to change "C" into "S" (for "substratum"):

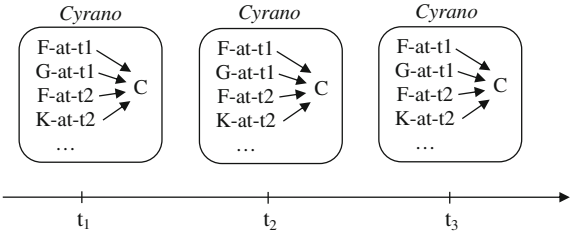


So I would like to suggest that it seems that the only difference we have here is terminological: in both cases, there is a unifying device included in the momentary temporal parts of Cyrano, and saying that in one case it's compresence and that in the other case it's a substratum does not make any difference in the way this unifying device manages to play its theoretical role. Let us now have a look at the endurantist alternative.

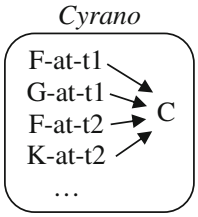
§7. Endurantism is the view that Cyrano persists through time and through intrinsic change by existing wholly at different times, rather than by having temporal parts. For the endurantist, one and the same (numerically identical) person exists wholly at t_1 and t_2 and has the two incompatible properties of having a big nose and having a small nose. In order to avoid the threat of having to deal with a contradiction, endurantists will typically embrace some kind of temporal indexation strategy and it is the standard version with time-indexed properties that I will be using here.[3] According to this view (*indexicalism*), Cyrano does not have incompatible properties, for instead of having properties like "having a big nose" he has time-indexed properties like "having-a-big-nose-at-t_1" and "having-a-small-nose-at-t_2", and these are perfectly well compatible and non-contradictory. Thus, since the indexicalist will claim that all properties are always indexed, no contradiction can ever arise from intrinsic change of an object that is numerically one and the same at different times, as the endurantist claims.

[3]I'll talk about adverbialism in §10.

Now, the BTT figure one is tempted to draw here is the following (as before, "C" is the compresence relation that ties together the different properties of the bundle):

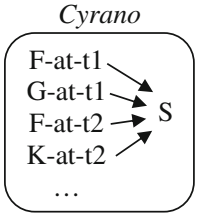


But this is a *bad* picture. It is *not* the picture the endurantist should be drawing. Firstly, why draw the same thing three times? These three bundles of properties are not three bundles located at three different temporal locations, rather they are one and the same. Drawing the same thing three times is here totally redundant and unhelpful. Secondly, the time-axis is also redundant—we do not need it in order to know what is happening to Cyrano at different times, since all temporal specifications are already included in the properties of which Cyrano is a bundle. So, after removing all of the redundant and confusing elements, here is the correct endurantist picture:



Cyrano is simply a bundle of time-indexed properties. If you find this strange and surprising … you really shouldn't. If you find it strange and surprising, it's perhaps just that you are accustomed to thinking in terms of perdurantism-inspired pictures. But I hope that you don't find this strange, and surprising—after all, this is exactly the endurantist picture Peter Van Inwagen draws (see Van Inwagen 1985, p. 195), except that he does not do it in terms of the bundle theory.

And again, with no surprise, the STT endurantist picture looks just like the preceding one:



The upshot of all the above considerations is, again, simply the following: the substratum and the compresence relation play the same theoretical role. Thus,

because both BTT and STT use their unifying device in the same way, they have exactly the same means to face the objection in a parallel way and it seems that the difference between them is merely terminological—one has a unifying device called "C" and the other has a device called "S" but since both devices are theoretical entities (they are there to do some theoretical work) and are thus individuated by their theoretical role, and since they play their theoretical role in the same way, they just seem to be one and the same thing under different guises. And if that's the case, there just does not seem to be any real difference between BTT and STT.

§8. I shall now more quickly go through the second of the two classic objections to BTT, which is the modal analogue of the first, and is structurally similar to it. Take Cyrano, who is a bundle of bundles of properties (or a bundle of time-indexed properties, if you prefer the endurantist approach). Now, the objector remarks, as before, that the identity of bundles depends on their constituents—a bundle must have the constituents it has, otherwise it would not be the same bundle. So, it seems that the components of a bundle are essential to it. But then, the bundle theorist faces the unwelcome consequence of his theory that any property of any individual turns out to be a necessary property of it. Take, again, Cyrano who has a big nose. In the bundle theorist's vocabulary, what we have is a bundle of bundles of properties, among which is the property of having a big nose. But, since Cyrano is this bundle, and since bundles have their components essentially, it is impossible for Cyrano to have had any other properties than he actually has—even the most insignificant and contingent ones, like the size of his nose, or the amount of hair he had this morning at 7 a.m. If this is true, bundle theory certainly does not look very appealing.

How does STT avoid this objection? It doesn't! The bare substratum cannot have different properties than the ones it actually has, since in itself it doesn't have any, and the thick particular (the substratum + the properties it bears) cannot have different properties either, *for exactly the same reason* that the BTT theorist's bundle cannot have different properties. If what individuates the particular Cyrano is *only* the substratum (which would be a strange view anyway) then it has all of its properties essentially, since it doesn't have any. If what individuates Cyrano is the substratum *and* its properties, then if you take one property away you don't have the same particular anymore, exactly as in the case of BTT.

Of course, the STT theorist, as before, will object: "I have something that you, BTT, don't have: a substratum that allows me to have a particular with different properties because it will still make it the same particular even if some properties change".

But if this were an acceptable reply here, then BTT has at hand exactly the same: "I have a compresence relation that remains the same even if it were to bundle different properties, and this guarantees me that the particular, while having different properties, is the same particular. You have your primitive unifying device that does the job, I have mine."

As before, it seems that both views can handle the objection in the same way by appealing to the unifying device, and that calling the device different names ("compresence" or "substratum") does not change anything since both unifying devices just play the same role in the same way. To be more precise, exactly as in the temporal case in the preceding section, both views in the modal case have the possibility to pick

their favourite view on persistence across possible worlds (trans-world identity, counterpart theory, modal perdurants, …) and use it to answer the objection. And as before, my point here is not to defend BTT against objections, but to show that whatever means STT has to avoid them can also be equally well used by BTT, simply because both views have a unifying device which just seems to be one and the same thing. As promised, methodology will be discussed later, in Chap. 4, but the general idea becomes easily apparent already: BTT and STT are 'equivalent' (in a sense to be specified), because the unifying device called "substratum" in STT and the unifying device called "compresence" in BTT are identical, because they play the same theoretical role in the same way, and they are theoretical entities (that is, they are individuated by their theoretical role)—more on this later.

§9. It is time now to compare the Bundle Theory with Universals (BTU) and the Substratum Theory with Universals (STU). As already mentioned, in my list of options in §2, (4) is ruled out for exactly the same reason (1) was: not all objects have the same number of properties, and so if one wants to have one and the same relation to be the bundling relation for all objects (which is plausible here since we are friends of universals) something like a variably polyadic relation (like "x, y, z, … are compresent with one another") is required.

But before, let's have a closer look at (5) where *different* universals play the role of the bundling relation for different objects; one universal per object. With respect to the two objections we have seen in the case of BTT and STT, the situation is here the same, and the same equivalence conclusion can be drawn. But there is another traditional objection to the bundle theory with universals that will perhaps make a difference between BTU and STU: the objection from the principle of Identity of Indiscernibles. Let us see how the objection goes.

BTU suffers from a traditionally weighty objection: it is committed to the principle of Identity of Indiscernibles. But this principle is false.[4] So, BTU is false.

$$[\text{Id.Ind.}] \quad (\forall x)(\forall y)((\forall F)(Fx \leftrightarrow Fy) \rightarrow (x = y))$$

Under BTU, material objects are said to be bundles of properties. Now, take two objects that have the same properties, for instance, as in Max Black's world, two perfect spheres of the same size, same mass, same composition, same colour, and so on. Both spheres are bundles of the same properties (universals[5])—and so are the same bundles. But then, the bundle theorist must accept that the two spheres are numerically identical—that is, there is only one sphere. And this is exactly what the principle of Identity of Indiscernibles claims.

But this principle is false, for it is quite possible there to be two numerically distinct objects that have exactly the same properties (that are qualitative duplicates). The example of *two* spheres exactly alike in all of their properties is possible.

[4]Or only *contingently* true.

[5]Of course, BTT does not suffer from this objection.

How can BTU get out of this trouble? Distinguishing between the two bundles by the use of spatio-temporal location properties ("being on the left of Cyrano") or by the use of haecceistic properties ("being identical to sphere A") has not proven to be a very appealing strategy in the abundant literature on this subject. But, for my present purposes, it is not the time now to examine these possible answers to the objection, rather what I wish to ask myself now is: how does STU face it? Or rather: why does this objection not even arise against STU? The answer is obvious and readily at hand: the two spheres are distinguished not by their properties (they *are* qualitative duplicates), but by what bears them, that is, the substrata that 'unify' them and put them together in order to make up an object. The substrata being numerically distinct, the two spheres are numerically distinct as well.

But what grounds the claim that the two substrata are numerically distinct? It cannot be a qualitative difference between them, so what is it? As we have already seen, there is not much of a choice, and so substratum theorists simply claim that the numerical difference between two substrata is a primitive fact. (Very well, any theory has its primitives.) But, as before, BTU can use exactly the same strategy—remember that here we have different compresence relations, one per object, and so two objects, even qualitatively identical, will always be numerically distinct since the bundling relation that ties together their properties will be a different universal—exactly as in the case of STU it will be a numerically different substratum. But then, as a *tu quoque*, one can ask: In virtue of what is a given compresence relation numerically distinct from another compresence relation? And there is no better answer to this question than to the same question about distinct substrata, the only option is primitive distinctness.[6]

But then, again, it seems that the thing that plays the role of a unifying device in STU (the substratum) and the thing that plays the role of a unifying device in BTU (the compresence relation) are both equally well suited to do the job. Once we are ready to accept as a suitable way to avoid the objection from Id.Ind. that there is primitive numerical difference between substrata, why not *as happily* avoid the objection by saying that there is primitive numerical difference between two compresence relations? After all, bundle theorists often speak about the compresence relation as of a primitive that is as under-defined and under-explained as a substratum is, so why not let it do the job for which it seems so naturally suited? So, as in the case of BTT and STT, the two theories here do have the same means to avoid any worries with Id.Ind., both contain a 'unifying device' that allows them to do so in the same way, and calling this device different names (substratum *versus* compresence relation) does not make a difference other than merely terminological. Stick to a neutral vocabulary (like "unifying device") and reformulate the two views in the light of the Id.Ind. objection: both will be able to say that sphere A and sphere B can be distinguished by there being a primitively distinguished unifying device for A and for B.

[6]The perhaps tempting idea to distinguish non-primitively between the compresence relations precisely by the number of properties they relate is unappealing because *some* distinct objects *do* have the same number of properties. (For instance two objects that are qualitative duplicates except for their colour.)

Note: In the same manner, BTU can avoid other objections that ultimately rest on the principle of Identity of Indiscernibles, like "the problem of angels" (Hawthorne and Cover 1998, p. 216) or "the problem with circular time and circular space" (Hawthorne and Cover 1998, p. 218)—indeed, these objections actually are 'just' different cases that exhibit the problem with identical indiscernibles in different ways. But, of course, Hawthorne and Cover appeal to a different version of BTU—the one that I shall examine in the following section.

§10. I will now consider places where the bundle theory and the substratum theory are *not* equivalent, and show that the reason why these versions are not equivalent is also the reason why they are much less appealing.[7]

A first such place is the last case (6) in my list of options from §2: BTU where the compresence relation is one and the same numerically identical variably polyadic universal that plays the role of the bundling relation for all objects (let's call it "BTU_2"). This feature makes BTU_2 more vulnerable to the objection from Id.Ind. and makes it a different, non equivalent, view than STU (and than the version of BTU examined in the preceding section (let's call this one BTU_1)). It is easy to see how and why: since BTU_2's unifying device is *one and the same* for all objects (rather than one per object as in STU or BTU_1), it is no wonder that we have troubles here with the identity of indiscernible objects, since the (allegedly) two indiscernible objects not only share all of their qualitative properties (universals) but they even share what makes them to be an object, they even share one and the very same unifying device! From this point of view, it is really not hard to see why worries arise about the claim that they are *two*, rather than *one*, in the first place! So, I dare say, what makes BTU_2 a different view from the others is also its main weakness.

John Hawthorne defends BTU_2 against this worry in his Hawthorne (1995), and his defence is as ingenious as it is simple: since, according to BTU_2, objects are bundles of universals, they can behave like universals; relevantly, a bundle of universals (for instance the bundle that is a sphere in Black's world) can be, exactly as a single universal can be, bi-instantiated, and bi-located. Black's world can be thus re-described in terms of BTU_2 as a world where there is one sphere bi-located at a distance from itself. (And this is strange, Hawthorne claims, only to the extent that the idea of a bi-located universal is perhaps strange.)

A consequence of this is that material objects, like a sphere or Cyrano or yourself, behave like universals. So such a view really seems to collapse the distinction between objects and properties—indeed, it seems that objects are simply eliminated from ontology. This is, in short, a worry put forward by Vallicella (1997). But it is hard to evaluate the dialectic force of this worry, since the BTU_2 theorist could very well simply bite the bullet, without perhaps too much harm. But Vallicella offers a

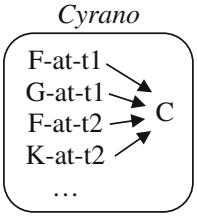
[7]Another such place, I believe (but without arguing for it here), is a bundle-theoretic-like view called the "nuclear theory", defended by Simons (1994) and based on Husserl's view. This view does not fit in my table from §2, indeed, it has a different structure than 'standard' bundle and substratum theories since, as Simons himself says, "it combines aspects of both bundle theory and substratum theory". I believe that precisely because of its unusual structure the view is unappealing, but I do not offer any arguments to support that belief here.

second objection that seems to me more damaging, while being simpler: to be multiply located, a bundle of universals would have to be instantiated, but this makes no sense for BTU_2. Here is an almost exact quote from Vallicella (1997, p. 94) that I have only slightly modified in order to stick to my terminology:

> A universal U is (multiply) located if and only if it is (multiply) instantiated. So if a bundle B of universals is itself a universal then it is (multiply) located if and only if it is (multiply) instantiated. But what could account for B's (multiple) instantiation? On BTU_2, universals are instantiated by being bundled together with other universals. But it makes no sense to suppose that B is bundled together with other universals; for B is a complete bundle of universals. […] But if B is not bundled together with other universals, then it is not instantiated. For on BTU_2, a universal is instantiated just in case it enters into a bundle. And if B is not instantiated, then it cannot be multiply instantiated. But if B cannot be multiply instantiated, it cannot be multiply located. So Black's world cannot be given Hawthorne's reading: it cannot be construed as a single sphere at a non-zero distance from itself. For the sphere cannot be doubly located without being doubly instantiated, and it cannot be instantiated at all, for the simple reason that a bundle of universals is not a universal but a particular, and no particular can be instantiated.

It is not my purpose here to try to refute BTU_2, even though I share Vallicella's worries. My point, as already mentioned, is simply to see that the reason why BTU_2 is a non-equivalent view, different from BTU_1 and STU, is also the reason why it is in trouble, where its competitors are on safe waters.

To make my case stronger, let me note another point of dissatisfaction with BTU_2 and the way it can handle persistence through time. Let us first suppose that our BTU_2 theorist is an endurantist. Remember the objection to endurantism from temporary intrinsics, and take again my neighbour Cyrano and say that at time t_1 he has a big nose, but he then decides to undergo plastic surgery (for expository reasons a bit later than before, say at t_5) and consequently has a small nose at a later time t_6. For the endurantist, this means that one and the same (numerically identical) person exists wholly at t_1 and t_6 and has the two incompatible properties of having a big nose and having a small nose. As we have seen, to avoid a contradiction, the endurantist appeals here to an indexicalist strategy and so this is how the view looked like:



Now, my worry is the following. At t_1, Cyrano has a big nose. At t_2, he has a big nose. At t_3, he still has a big nose. And so on, until the surgery. The intuitive thing to say here is clearly that Cyrano keeps having a certain property for a certain time—but the indexicalist endurantist just cannot allow for that. According to indexicalism, at any time during the interval t_1–t_4, Cyrano has to lose all of his properties and gain new ones: he first has the property "having-a-big-nose-at-t_1", then the property "having-a-big-nose-at-t_2", then the property "having-a-big-nose-at-t_3", and so on.

According to this view, because Cyrano cannot simply (*simpliciter*) have the property of having a big nose, he has to change his properties all the time, and he cannot keep any—he just cannot stay the same. And since the property "having a big nose" is not available to her, the endurantist does not have the theoretical means to say that all these time-indexed properties have 'something in common'—they just are different properties.

But rather than objecting to endurantism, my point here is that time-indexed properties *are tropes*. In the indexicalist's world there is no room for a single property to be multiply instantiated, since any property is time-bound and cannot be instantiated at different times, and so there simply is no room for universals (multiply instantiable properties).

And of course, not only properties have to be time-bound but space-time-bound. The need for this is most salient in the case of a time-travel scenario. Suppose that at t_6 Cyrano travels back to the past in order to tell his former self that the surgery will be all right and that he does not have to worry. According to endurantism, Cyrano then has the properties "having-a-big-nose-at-t_1" and "having-a-small-nose-at-t_1": a seeming contradiction, easily solved by claiming that all properties are always space-time-bound, since of course "having-a-big-nose-at-l_1-t_1" and "having-a-big-nose-at-l_2-t_1" are not contradictory (where "l" stands of course for "spatial location").[8]

So: the endurantist has to do something in order to avoid the Lewisian worry about temporary intrinsics, and if what she does there is to embrace indexicalism, her properties just have to be space-time bound and non-multiply instantiated—tropes. So it seems that this is not an option the BTU_2 theorist can choose.

But perhaps she has other options—she can either choose (to try) to be a perdurantist or (to try) to remain an endurantist but abandon indexicalism in favour of adverbialism. Let us examine these two options in turn.

Perdurantism just does not seem to be available to the BTU_2 theorist. At the very least, it would be very strange for her to take that route since her central claim is that objects behave like universals and can be multiply located, while the perdurantist's central claim is that all objects are space-time bound and that nothing (no object) can be multiply located. So even if perdurantism does not force one to embrace tropes (unlike indexicalist endurantism), it does not seem to be a viable option for the friend of BTU_2.

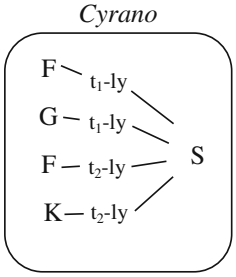
The last option is endurantist *adverbialism*. Or is it? Indeed, we will now easily see that adverbialism is available only to the *substratum* theorist, and not to the bundle theorist and so this strategy cannot be of any help to BTU_2 either.

The adverbialist solution to the problem of temporary intrinsic properties proposes to temporal modify not the property but the having of it. Thus the adverbialist will say that "Cyrano has a big nose at t_1" is to be analyzed as "Cyrano has-at-t_1 a big nose" or, more elegantly, "Cyrano has t_1-ly a big nose" (see Johnston 1987, p. 129). So, according to adverbialism, there is not just the having of a property,

[8] I am not saying that this solves all problems the endurantist might have with time travel.

there is always t-ly having (or having-at-t) of a property. This will provide a solution to the problem of temporary intrinsics, because Cyrano has a big nose at t_1 and has a small nose at t_2, and so he has both the incompatible properties, but he has the former t_1-ly and the latter t_2-ly and this is how the threat of a contradiction is avoided.

Now, in order to be able to be an adverbialist, the *substratum* theorist might want to say that there is a *third* component in her view—a relation of exemplification between the substratum and its properties—and that this relation is time-indexed (rather than the properties being time-indexed, as the endurantist indexicalist would have it). The endurantist-adverbialist-substratist picture then looks like this (where "S" stands for "substratum"):



The bundle theorist, obviously, cannot (and does not want to) provide anything like this, since she does not introduce a substratum that needs to be related by a special relation to its properties; her view does not require any such intermediaries—and so, she cannot be an adverbialist since there is no suitable place to put the adverbialist index,[9] and so BTU_2 simply cannot appeal to this strategy. To sum up:

First, BTU_2 is not compatible with endurantist indexicalism because this view requires tropes rather than universals; it does not, at least *prima facie*, look like it could be made to work under perdurantism; and endurantist adverbialism simply is not an available option. Add to this Vallicella's worries, and it really seems that BTU_2 is a non-equivalent view to the others only because of features that make it ultimately very hard to sound appealing.

Second, there is another place where the bundle theory and the substratum theory are not equivalent: only the substratum theory is compatible with adverbialist endurantism. But, this is true only for a certain version of the substratum theory, namely a version that insists on there really being a *third* component in the theory, a (time-indexed) *relation* between the substratum and its properties. And even substratum theorists themselves often agree that this is a bad version of their view (among other reasons, because of Bradley-like regresses and related issues). Most recently, Sider (2006) in his defence of substrata insists that we should not put too much weight on the relation of exemplification. It is often claimed that exemplification is not a relation, that it is a "non-relational tie", and that we

[9]Indexing the relation of compresence would make it a *perdurantist* view.

shouldn't reify exemplification (Lewis 1983, pp. 351–355). I guess that the idea here is to build already into the substratum the theoretical function to be able to simply stick to its properties in addition to its function of being a unifier (exactly as compresence!), rather than only taking it as a unifier that requires a further theoretical device in order to account for the nature of objects. And only if one takes the latter (much less palatable) option, will one get a different view than the bundle theory (and here again, the reason why it is different is also the reason why it is unpalatable).

§11. Until now, I have spent some time focusing on the bundle theory and the substratum theory, in order to have a very detailed example of a case of two theories that are (in all except the unpalatable places) equivalent—in a sense of "equivalent" still to be discussed explicitly. This debate is, however, not an isolated case. Indeed, throughout the history of philosophy, several other debates have been modelled as oppositions between a 'bundling' approach and a 'substratist' or 'substantivalist' approach. Such debates obviously include the broadly Humean bundle theory of the Self, and its alleged opponent in the form of a broadly Cartesian substance theory of the Self—these are, in fact, 'only' particular cases of the general bundle theory and substratum theory I have discussed above. More interestingly, the controversy between *relationism and substantivalism about time* is also a very similar case. In all these debates, the substantivalist side typically insists that in order to provide a good treatment of the subject-matter of the theory (time, Self, material objects), it is necessary to postulate the existence of a certain kind of substratum or substance[10] that is required to account for some important issues such as particularity, individuation, unity, independence, persistence,… and that allows in this way to solve some puzzle cases, intriguing phenomena, or philosophical problems. Without an underlying substance, the friends of the substratum theory feel that properties would go 'floating free' and objects like tables would lack particularity, the Self would lack unity, and nothing could genuinely persist through time while undergoing intrinsic change. Substantivalists about time would analogously feel that events and things need to 'occur at' or 'be located at' times that need to be substantial enough to be able to 'support' them or 'contain' them.

The other side, the relationist one, typically feels that this is an unnecessary expense and that one can get the job done in an ontologically cheaper way. Relationists will thus deny the need for any substance, they will claim that it is a mysterious thing that we should spare ourselves, and that it is enough, to account for all phenomena, puzzle cases, and philosophical problems, to use the entities that substantivalists take to be 'had' or 'united' by the substance (properties, events, …) and explain how these are inter-related. According to relationists, who often declare themselves to be more respectful of Occam's Razor, it is enough to have properties tied together by a special relation of 'compresence' to get ordinary objects or Selves

[10]The word "substance" is a tricky one, and it often means very different things in the mouths of different philosophers. I shall use it as a synonym of "substratum", and when it matters (see below) I will say precisely what it refers to.

(that are bundles of properties) and it is enough to have events related by a special relation of simultaneity to get times (simultaneity classes of events).

In the case of the general bundle theory and substratum theory, we have seen that the alleged disagreement is actually minimal. Now, it is interesting to see not only that the other two debates (about the Self, and about time) are similar in nature, but also that we step on very similar arguments and objections in all three cases. To mention now only one, the problem with Identity of Indiscernibles I have discussed above parallels the problem often raised against the relationist theory of time: the case of 'Time without Change'. Both problems yield the same difficulty, except that in one case the question is about how to account for numerical diversity of objects that have the same properties (that are qualitative duplicates) and in the other case the question is about how to account for numerical diversity of times (instants) that 'have' the same events, that is, at which the very same events occur during a period of time without change.

In the following sections, I shall now continue to elaborate my claim that there is much less of a disagreement between relational ontologies and substantival ontologies than it is usually thought. As we have already seen in the case of the bundle theory and the substratum theory, and as we will now see in the case of relationism and substantivalism about time, the two sides of the debate are not very different from each other. We will see again that both the relational side and the substantivalist side work in the same way, suffer from and answer the same objections, and are structurally extremely similar. Of course, a genuinely strong claim of metaphysical equivalence would require a very detailed examination of the many variants there are of relationism and substantivalism, in the way I did it above for the bundle theory and the substratum theory. I will not attempt to elaborate such a detailed claim here. What I will do is to point to certain structural and dialectical features of the two sides of the debate that are common to most variants of relationism and substantivalism, and make a claim that will be detailed enough to see how an equivalence claim between the two views can be construed.

§12. A useful and standard way to introduce the substantivalist theory of time is by the use of 'the container' metaphor and the two central arguments in its favour: the possibility of 'empty time' and the possibility of 'time without change'. According to substantivalism, time is like a container in which events and things are placed, a container that exists independently of what is placed in it. While I am typing this sentence, the container is not empty but, importantly, it might very well be: a container is perfectly capable of not containing anything. Less metaphorically, time is a substance that exists independently of events and things located in time, and consequently it is such that it allows straightforwardly for the possibility of there being periods of time during which time continues to pass even if no changes occur (so that the universe is 'frozen' during this period of time) or even if nothing at all occupies it (so that time is 'empty' during this period).

Contra substantivalism, the relationist theory of time rejects the idea of time as being independent of events and things placed in it. Rather, relationists claim, time is nothing over and above temporal relations among events and things located in it. Thus formulated, relationism probably sounds too circular, so let us try to put it in a

different way: if there were no objects and events, there would be no time, for time is not a thing (a substance) but rather a system of relations among events and things. A particular instant of time is thus, according to relationism, a collection of simultaneous events and things (a simultaneity class of events and things), and a time-series is all the collections of simultaneous events in the order in which they occur. It is at least a *prima facie* consequence of this view that it cannot accommodate the idea cherished by substantivalists that there could be 'empty' periods of time or periods of time without change.

The latter claim is something that will be one of the main points of discussion below. But before I start, let me make some very quick terminological remarks: substantivalism is also often referred to as 'absolutism' (following Newton's absolute space and time theory) or 'platonism' (since Plato was among its prominent defenders). The term 'substance' must here be clearly distinguished from Aristotle's use of it. The Aristotelian theory of substance is something different and is a theory about the nature of ordinary material objects. If any theory of material objects is analogous to substantivalism about time it is not the (Aristotelian) substance theory but rather the substratum theory. Relationism is also often referred to as 'reductionism about time' for obvious reasons.

As metaphysical theories go, it seems at a first sight, and probably even at a second deeper look, that these two rival views could hardly be more dissimilar and opposite to each other. To put it in terms of the recent debate in meta-ontology, the dispute between the substantivalist and relationist theories of time seems to be a good candidate for a clearly substantive non-verbal and non-trivial one. But, as in the case of the bundle theory and the substratum theory, we will see the two allegedly very different rival views are much less different than has been thought: their structure is extremely similar, their strategies are extremely similar, they can both face the possibilities of 'empty time' and 'time without change' *in the same way*, so that, as we will see, some central objections to one side always have a sneaky tendency to reappear for the other side as well. In the face of this, I will then put forward two possible conclusions that can be drawn (here, some of the methodology will already become apparent): either a strong meta-metaphysical claim that these two views turn out to be 'equivalent' or a weaker claim that the two views are so similar and work in such similar ways for all theoretical purposes that there is little reason for choosing one rather than the other.

§13. Let us start with the case of the possibility of time without change that nourishes a large debate between substantivalists and relationists. Let me quickly summarize the well-known Shoemaker (1969) argument that intends to show that such periods of time when all changes in the universe come to a stop, called 'global freezes', are indeed possible. It is not my purpose, as I shall explain below, to defend or reject this argument (rather, I will be interested in its conclusion and its implications for the substantivalism-relationism debate), but it will be helpful to bear it in mind. The purpose of the argument is to show that in a possible world where 'local freezes' occur, its inhabitants can have a good reason to think that 'global freezes' occur (even if, of course, no one can directly experience them). Take a world W divided into three spatial zones A, B, and C. There are local

freezes: at regular intervals, in each of the zones all changes come to a stop for a certain period of time, while at least one of the other two zones remains unfrozen. This happens for one hour every 2 years in zone A, for one hour every 3 years in zone B, and for one hour every 5 years in zone C. Thus, the inhabitants of W who can be aware of local freezes when they occur in a different zone than the one they are located in, and who have made the calculation, have a good reason to believe that every 30 years there is a one hour global freeze. To make their reasoning stronger, it can be added into the example that in every zone, just one minute before a local freeze occurs some visible changes occur to 'announce' the freeze—for instance, just before a local freeze occurs all things located in the zone in question turn red. Every 30 years, it is then not only a simpler theory (it is simpler to say that local freezes occur with a regularity rather than to say that there is an exception in the regularity every 30 years) but also the fact that everything, in all zones, turns red one minute before the expected global freeze, that indicate that indeed such a global freeze is about to take place. It is thus, in W, reasonable to believe that there are regular one hour periods where time continues to flow while no change at all occurs.

This argument actually does not *show* that it is possible that there can be time without change, since local freezes are simply presupposed without argument. What is interesting, then, about this argument is not so much what it shows, but that it provides a useful metaphysical scenario that has traditionally been taken to have important implications with respect to the debate about the nature of time. Indeed, if global freezes are possible (or if we could have good reasons to think that they are possible) then this would show that substantivalism has to be true (since relationism construes time out of changes), whereas if such global freezes were shown to be *im*possible this would leave room for both substantivalism and relationism to be true.

What I intend to do now is to show that in *both* cases, *both* theories can equally well do the job. Just as it would be a mistake to think that the *im*possibility of global freezes shows that relationism is true (since substantivalism can accommodate this possibility as well), it is also a mistake to believe that the possibility of global freezes shows that relationism is false, since as we shall see it can accommodate this possibility *in the same way* substantivalism does. My strategy is thus completely different from the strategy that modifies ('modalizes') relationism by taking other-worldly changes as being constitutive of this-worldly instants (see Forbes 1993); rather I shall show that relationism *as it stands*, without any modification of the theory, can actually deal with the possibility of time without change in the same way substantivalism does.

Let us start by examining the substantivalist's strategy more closely. How exactly does *substantivalism* manage to accommodate the possibility of time without change? Suppose that there is a global freeze and that there is no change going on—how can time continue to flow? The question is, what makes it true that there is a series of non-identical instants one after another, rather than just one single instant? How can the instants in this series (the series of instants that occurs during a global freeze) be distinguished? Well, of course *not* by what

changes/events occur at them, since these are all the same. From the *qualitative* point of view, they are indiscernible—if this were the criterion for distinguishing instants, then there would only be one instant, and no global freeze that lasts for an hour. So what makes different instants different, if they are not qualitatively different? The answer is: they are primitively numerically distinct; they do not need to be qualitatively discernible, since instants in themselves do not have a nature such that they are capable of being qualitatively distinct one from each other; rather they are fundamental, non-decomposable, primitively numerically distinct entities. Following the container metaphor, during global freezes there is a series of containers that in themselves are qualitatively indistinguishable and that contain qualitatively indistinguishable stuff, but that are primitively numerically distinct.

This is what I will call a 'problem-solver' (and this is where some bits of my overall methodology leak out). Simply put, a problem-solver is something that is a primitive in a theory and that solves a problem. Perhaps, every primitive in every theory is a problem-solver—for why do we introduce primitives in the first place, if not for them to do an explanatory job? And how do they do this explanatory job? By having a primitive capacity to do so. I shall give some more examples below, but for now let us stick with substantivalism and the problem of time without change. How can the theory account for there being global freezes? By having numerically distinct instants that are not distinguished qualitatively, but primitively. The premise that there are primitively numerically distinct instants is thus a 'problem-solver' in the sense that, without it, the theory would not be able to face the scenario of time without change, and that it succeeds to do so only in virtue of the postulation that it can do so. The latter claim may sound a bit pejorative with respect to substantivalism, but it is not: every theory has its primitives and every primitive is, at least to some extent, a problem-solver. As I see it, the use of problem-solvers is commonplace in all philosophy, and without it we would not get very far—it just is one among the components of the philosopher's toolkit.

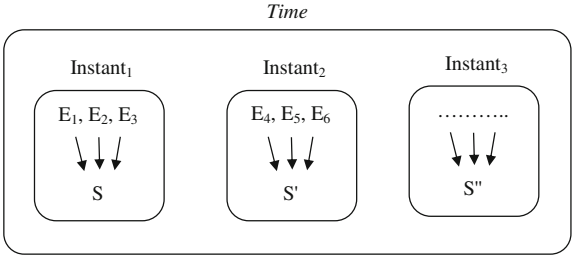
Let us now turn our attention to relationism. According to this view, an instant is a simultaneity class of events, more precisely, it is a bundle[11] of events that are put together by the relation of simultaneity, and so it is individuated by these events and this relation. An instant thus has a qualitative nature, unlike under substantivalism, and instants can be in this way distinguished by the events they contain. But when a global freeze occurs, all of the instants during this one hour period contain the same events, and so are indiscernible; consequently they cannot be qualitatively distinguished any more, and, the objector claims, one has to conclude that there actually is only one instant—and so such a view cannot accommodate the possibility of a one hour global freeze.

I think that the relationist has a reply readily available at hand here. Consider more closely what the relation of simultaneity is and what it does. Its theoretical role is such that it is a function that takes events as input and gives an instant as output. For each instant there is such a relation, and this relation is not and cannot

[11] *Not* a set, since instant are not abstract entities, on this view.

be one and the very same relation for all instants—otherwise, regardless of there being global freezes, there would be no more than one single instant. (If properties and relations are tropes, the situation is even clearer: tropes being non-repeatable entities it would not even be possible for the relation of simultaneity to be one and the same for different instants, so what we have here are exactly resembling and numerically distinct tropes of simultaneity, one per instant. If properties and relations are universals, there are two *prima facie* possibilities: either the relation of simultaneity is one and the very same relation for all instants, or it is a different universal for each instant. As suggested, the former possibility yields difficulties even if no global freeze occurs—the case of a global freeze is just the most salient case where these difficulties become apparent—so it is the latter that should be endorsed anyway. Alternatively, claiming that there are numerically different *instances* of *one* universal of simultaneity could perhaps also do the job.[12] What is important for me here is that there is always something numerically different for each instant that is responsible for tying up together the events to make up the instant.)

Thus the structure of the relationist theory of time is the following:



Each instant is made out of events tied together by a relation of simultaneity that is different from one instant to another. Time is then a series of such instants.

Now, what exactly happens when a global freeze occurs? The events that compose the various instants that occur during the freeze are the same, since no changes occur. But that does *not* at all prevent relationism from accommodating the claim that there is a series of numerically different instants: a series of instants that lasts one hour where each instant contains the same events E_1, E_2, … E_n *but* tied together by a different relation of simultaneity, as is anyway the case even when no freezes take place. The instants will thus be distinguished not qualitatively but numerically by the relation of simultaneity that individuates them as well as the events do. And how is the relation of simultaneity distinguished from one instant to

[12]This is similar to a strategy that Paul (forthcoming_a) explores with respect to the Bundle Theory of objects, when she says: '[…] properties are shared, while property instances are primitively individuated. On this approach the explanation of the possibility of the qualitative indiscernibility of the spheres in W is based on an underlying identity of properties, while the numerical difference between the spheres reductively supervenes upon the numerical difference of the property instances in each bundle'.

another? The question is: how is it (and not *why* is it, we have seen why above) that the relation of simultaneity for an instant is distinguished from the relation of simultaneity for another? How is it not one and the same relation? As in the case of substantivalism, the answer is primitivist: it is primitively the case that the various relations of simultaneity can be said to be numerically distinct (but of course not qualitatively distinct) one from each other.[13] As before, this is a 'problem-solver'.

Furthermore, it is a problem-solver that strictly parallels the one that substantivalism uses: both views can only face the 'time without change' scenario by using primitive machinery that distinguishes numerically different instants during a global freeze. Functionally, both problem-solvers used in both views are, with respect to the problem of time without change, equivalent. One side calls it a 'substance' (a 'container', a 'time') and the other calls it a 'relation of simultaneity'; but when you look at what they are *doing* in the theory that employs them, it is actually really hard to tell them apart. In other words, in one case the problem-solver is such that events are said to be 'placed' in it, or 'contained' in it; and in the other case the problem-solver is such that events are said to be 'tied together' by it—but, such metaphors aside, the functional role of these problem-solvers, with respect to the problem of time without change, is the same. Both have the primitive function of making different instants (numerically) different, and both can thus equally well do the job of accommodating the possibility of global freezes. And it is no wonder that they can both do the job since they are primitives and any primitive can be given any power one wants to give it, especially if one's opponent in the debate does the same. What I have in mind here is a view about the nature of primitives in metaphysics, such as the problem-solvers involved here, that takes very seriously the *functional role* they play in the theory. A primitive being primitive, it is non-analysable by its very nature. We are not really given any information concerning its nature; we are told *what it does* rather than *what it is*. So it is what it does that counts—after all, that's what any primitive is introduced for in a theory in the first place (otherwise there would be little justification for having it). Thus, primitives are individuated by what they do, what their functional role in a theory is, and as a consequence two primitives that do the same job just turn out to be equivalent, for all theoretical purposes. I have shown above that with respect to the case of time without change, the relationist and the substantivalist primitive machinery does the same job at the same place in the same way (that is, in a primitive way). Now, this does not mean that the two theories themselves are equivalent, since there may be other places where they are different. I shall now examine the case of 'empty time' that may perhaps be such a place.

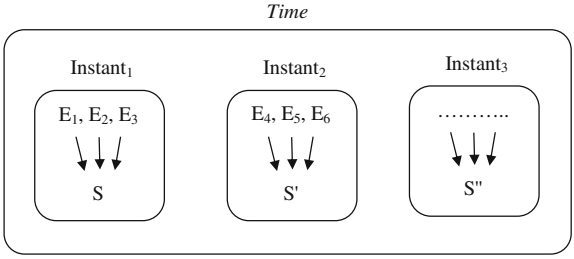
§14. Up to now, we have seen the case of relationism and substantivalism with respect to the problem of time without change, and we have seen that both views behave in fundamentally the same way. In this final section, I shall now consider

[13]To illustrate this point, take the case where properties and relations are tropes—in this case, what I say parallels the claim that one trope of simultaneity is exactly similar to but primitively numerically different from another trope of simultaneity.

the case of 'empty time' that is different from the case of time without change (an empty time or an empty period of time is a time at which nothing instantiates any properties, there just does not exist anything at such a time or period of time, while in the case of time without change things and/or events do exist, it's just that they are the same at different instants during a period of time without change). Prima facie, we have here again a place where relationism and substantivalism do not behave in the same way (and that also parallels a similar case in the bundle theory and substratum theory debate, as we shall see). But, although the conclusion will be less straightforward than before, I will also argue that, at the end of the day, both relationism and substantivalism can also treat the problem of 'empty time' in the same way.

Look again at the schema in the previous section, to quickly remember the structure of the relationist theory, in the normal case when times are not empty. On this schema 'S' stands for the relation of simultaneity, a different one per instant. Now, the objection goes, such a theory cannot accommodate the possibility of there being 'empty time', that is, of there being a series of instants at which no events occur (as opposed to the problem of time without change where the challenge was to accommodate the possibility of the same events occurring at different instants). Indeed, it seems that relationism just needs some events to be there, since events are constitutive of what instants are.

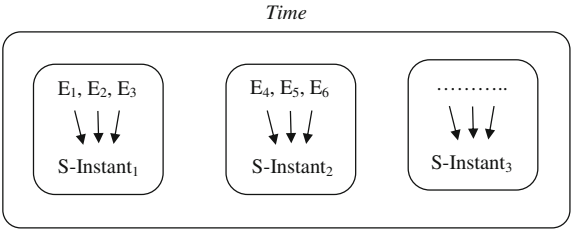
Before I go any further, let me first ask: how does *substantivalism* accommodate the possibility of empty time? There are two possibilities, one of them being certainly more natural for a substantivalist to embrace than the other. To distinguish them, let us see what the substantivalist view looks like in the normal case when time is not empty, and let us quickly examine the not-so-natural option first. According to this version of substantivalism, instants are 'thick', that is, the substantivalist's picture is the following:



When conceived of as 'thick', instants are such that they are made of a substance and of the events that are had by this substance—thus, instants are not just the substance, they are the whole. If instants were 'thick' in this way, the picture would actually look exactly as the relationist one. On purpose, I have chosen the same letter 'S' to stand here for 'substance-instant', in order to make it clear that if this were the substantivalist conception, there would really be no difference with respect to the structure of the theory between this view and relationism. Instants would be construed out of S and events, and calling S different names ('substance' for the

substantivalist or 'simultaneity' for the relationist) would hardly make any difference other than a terminological one.

But, most likely, substantivalists would not be happy with such a picture of their view, rather it would be more natural for substantivalism to see instants as 'thin', in the following way:



According to this picture, instants are not made out of events, they are the 'Ss' of the pictures above. If the substantivalist wants to, she can speak of 'thick' instants, of course (that is, S-Instant$_n$ + E_k, E_l, E_m, …), but this is not her primitive and fundamental notion of an instant. Rather she sees instants as being thin and ontologically independent of events, and this is how her view can easily and straightforwardly accommodate the possibility of there being a series of instants at which no events occur at all, that is, the possibility of empty time (so that in the figure above there would be no 'E_1, E_2, E_3, E_4, …' events, there would be just S-Instants). Of course, as I have already said when discussing the problem of time without change, substantivalism can accommodate this possibility only with the central help of its now familiar problem-solver: primitively numerically distinct instants. Indeed, the various instants included in the series that forms an interval of empty time cannot of course be distinguished qualitatively, since they are thin and have no qualitative nature at all, and since no events occur at them, so they have to be distinguished primitively (numerically).

The substantivalist has to defend the plausibility of such a view, exactly as the substratum theorist has to defend the plausibility of there being propertyless substrata, supposing that she wishes to do so, instead of thinking of substrata either as thick (substratum + properties) or as being thin but necessarily such that they exemplify some properties. The latter possibility is also relevant for the substantivalist: if she wishes to accommodate the possibility of empty time, she must not only defend the view that instants are not made out of events, but also the stronger view that instants are not necessarily such that some events occur at them. Suppose she can do this. Actually, she very easily can. The notion of an instant (as the notion of a substratum) is her primitive; it is a primitive postulate of her theory to which she can give any powers she likes: this is what primitive functions in any theory are for, and this one can be a function that can play the role of a lonely, empty instant or the role of an instant that necessarily contains some events—this all depends on whether one wants to accommodate the possibility of empty time or not. It works

likes this: *first*, it has to be decided, for independent reasons, whether it is a good thing or not to accommodate the possibility of empty time, and *then* it will be decided whether such a function is incorporated in the notion of a substantivalist's instant or not, and it can easily be said to be both ways. Instants are, after all, primitive theoretical postulates and problem-solvers, and one can simply define them to be one way or the other—the important question being thus not the one about what the theory can or cannot accommodate but about what we want it to be able to accommodate. The powers of our primitives are entirely in our power; they are problem-solvers that are not defined by what they are (since, being primitives, their nature is unanalyzable so we don't know much about what they are), but by what they do, that is, by what function they play in a theory, and this is something that is up to the theorist to decide.

Now, my point is that if that's what substantivalism can do, relationism can do it as well. Suppose that the relationist wishes to accommodate the possibility of empty time. She, too, has a problem-solver in her theory that can, as we have seen above, account for numerical difference between instants in the case of time without change. This part of the problem—accounting for numerical diversity of instants that form a series that is an interval of empty time—is then easily done in the same way. Now the second part of the problem remains, which is the capacity of the theory's problem-solver to exist 'alone' without any events. Granted, it does sound better to say that a substance like a substantivalist's instant can stand alone without there being any events than to say that a relation like the relationist's relation of simultaneity can stand alone without any events as its relata. But, as I have suggested above, these labels, like 'substance', 'substratum', 'simultaneity', and 'compresence' are no more than useful metaphors in the same way 'a container' was a useful metaphor to introduce the substantivalist theory. These metaphors help us understand better what is being said to us, and that is something important of course. But what I want to emphasize is that at bottom the referents of these metaphors are functional primitives postulated by a theorist and needed for her theory to work, they are problem-solvers whose nature is not analyzed but stipulated, and whose role in the theory is what counts. If it makes sense to claim that an instant can be empty, like a container can be empty, if it makes sense to claim that a substratum can exist without exemplifying any properties, then it can also make sense to claim that the relation of simultaneity can tie no events at all, or that the relation of compresence can tie no properties at all, perhaps in a similar way one would construe an empty set.

This parallels a general objection raised against the bundle theory, namely that while substrata are ontologically independent, the relation of compresence, as with all relations, is dependent on its relata. 'Independent' here probably means that it can exist independently of the other elements (properties)—and bear in mind that a strong version of this claim is needed, namely not just that the substratum can exist independently of this and this property but that it can exist 'alone' independently of having any properties at all. It is often said that substrata can satisfy such a requirement, while the relation of compresence can't because relations and properties cannot 'float free'. But this is no more than a familiar prejudice against the

bundle theory. What I call prejudice here, Hawthorne and Cover call simply 'incredulous stare' (while speaking about the bundle theory combined with universals):

> Perhaps some philosophers will claim to find it just self-evident that universals are had by something. We don't have much to say to such philosophers. We do note, however, that the polemic against the bundle theory has rarely taken the form 'It is simply self-evident that anything quality-like is directly or indirectly predicated of something that isn't like a quality […]'. If opponents of […] the Bundle Theory wish to retreat to this form of an incredulous stare, so be it. (Hawthorne and Cover 1998, p. 207)

Yet a different way to address this issue can be found in Gallen Strawson's article on the substance theory and the bundle theory of the Self:

> 'But if there is a process, there must be something – an object or substance – in which it goes on. If something happens, there must be something to which it happens, something which is not just the happening itself'. This expresses our ordinary understanding of things, but physicists are increasingly content with the view that physical reality is itself a kind of pure process – even if it remains hard to know exactly what this idea amounts to. The view that there is some ultimate stuff to which things happen has increasingly ceded to the idea that the existence of anything worthy of the name 'ultimate stuff' consists in the existence of fields of energy – consists, in other words, in the existence of a kind of pure process which is not usefully thought of as something which is happening to a thing distinct from it. (Strawson 1997, p. 427)

This being said, let me come back to my preferred way of addressing this issue, to insist that our metaphors do play an important role here. Here is how I think that it works: first, for intuitive reasons or for independent philosophical reasons, a theorist more-or-less explicitly decides whether or not it is desirable for her theory to allow for the possibility of empty time; second, she has to decide how to accommodate it and does this by incorporating in her primitive problem-solver the power to do so; and third, she has to make her problem-solver graspable by others and express it in a way that conveys well the concept she has in mind—here the metaphors play an important role, since by calling her problem-solver a 'substance' the theorist probably better conveys the idea that time is independent of events and, consequently, that there can be time without events, than if she calls it 'a relation of simultaneity'; although, as we have seen, both options are no more than different ways of expressing oneself and both can equally well accommodate the possibility of empty time. The choice between substantivalism and relationism, with respect to the problem of time without change and the problem of empty time, is thus a choice between alternative ways of formulating the same thing, where nothing really depends on the formulation except that, of course, one formulation can be better than another at expressing what one wants to say in a more understandable way.

§15. This completes my discussion of the claim that the pairs of views we have seen above (the bundle theory and the substratum theory, relationism and substantivalism) are not as different from each other as what we could have thought. In fact, they appear to be able to face the same theoretical challenges in the same way—mostly, in a primitivist way. In Chap. 4, I shall return to these issues and explicitly

discuss what kinds of conclusions of metaphysical equivalence one can draw from this. But before I do so, I will first discuss in detail cases of theories which appear to be only partially functionally equivalent (in Chap. 2), and theories which are not functionally equivalent at all (in Chap. 3).

Chapter 2
Partially Equivalent Metaphysical Theories

Abstract In this chapter, I discuss the case of a metaphysical debate that has been the target and centre of interest for many of those who work on meta-metaphysics, namely the problem of how objects persist through time: the endurantism versus perdurantism controversy. Some have argued, for various reasons, that this debate is a good example of a merely verbal one, where two allegedly competing views are in fact translatable one into the other—they end up, contrary to appearances, to be equivalent. In my discussion, I conclude that this is correct, but only to some extent, and that there does remain room for substantive disagreement. The second thing that I wish to achieve in this chapter is to start to defend a metaontological view that emphasizes a point which I think is often taken and acknowledged by many of those who are involved in metaontology, but which is not so often explicitly defended, namely that when asking the question "Are metaphysical debates substantive or verbal?" the correct answer is "It depends." Some debates are substantive, some debates are merely verbal, sometimes it is true that a problem or a question can be formulated in equally good frameworks where there is no fact of the matter as to which one is correct or where we just cannot know it. Furthermore, importantly, as my examination of the persistence debate will show, there is room for the view that such a debate is largely merely verbal but not entirely and that some parts of it are substantive, and decidable by philosophical methods. It is possible, and it is the case with respect to the persistence debate, that inside a debate some points are merely verbal while other are places of substantive disagreement. A moral of this is that, at the end of the day, the best way to do meta-metaphysics is to do first-level metaphysics.

§1. In this chapter, I shall have a close look at one metaphysical debate that has been the target and centre of interest for many of those who work on meta-metaphysics, namely the problem of how objects persist through time: the endurantism versus perdurantism controversy. McCall and Lowe (2003), Miller (2005a), and Hirsch (2008) have all argued, for different reasons and in different ways, that this debate is a good example of a merely verbal one, where two allegedly competing views are in fact translatable one into the other—they end up, contrary to appearances, to be equivalent. In my closer look at this debate, I will conclude that this is correct, *but*

J. Benovsky, *Meta-metaphysics*, Synthese Library 374,
DOI 10.1007/978-3-319-25334-3_2

only to some extent, and that there does remain room for substantive disagreement. To do this, I shall proceed somewhat differently: instead of looking for a general way to translate or to make equivalent the two (actually, more, as we shall see) competing views, I will go through several *first-level metaphysics* steps and look for places where alleged disagreement turns out to be merely verbal.

The second thing that I wish to achieve in this chapter, that was already apparent in Chap. 1 as well, and that will reveal a bit more about my methodology, is to exhibit a metaontological view that emphasizes a point which I think is often taken and acknowledged by many of those who are involved in metaontology, but which is not so often explicitly defended,[1] namely, that when asking the question "Are metaphysical debates substantive or verbal?" the correct answer is "It depends." *Some* debates are substantive, *some* debates are merely verbal, *sometimes* it is true that a problem or a question can be formulated in equally good frameworks where there is no fact of the matter as to which one is correct or where we just cannot know it. Furthermore, importantly, as my examination of the persistence debate will show, there is room for the view that such a debate is *largely* merely verbal but *not* entirely and that some parts of it are substantive, and decidable by philosophical methods. It is possible, and it is the case with respect to the persistence debate, that *inside* a debate some points are merely verbal while other are places of substantive disagreement. A moral of this is that, at the end of the day, the best way to do meta-metaphysics is to do first-level metaphysics, from which meta-metaphysical claims (such as equivalence claims) can arise. The priority should be given to the low-level considerations, and meta-metaphysical claims should not be made in too general a way but should come from particular decisions taken case by case on the level of metaphysics.

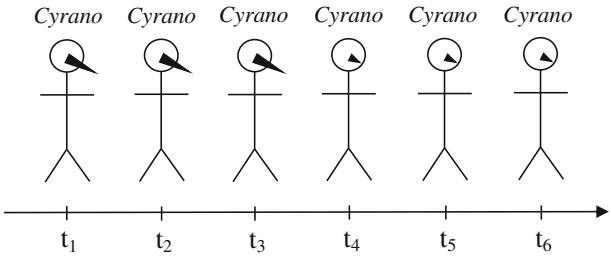
§2. In this chapter, I will focus on perdurantism and endurantism under the assumption that eternalism is true. Presentism (and eternalism as well) will be discussed in Chap. 3. Perdurantism comes in two main versions—the *worm view* and the *stage view*—and endurantism comes in two main versions as well—*indexicalism* and *adverbialism*. I will now carefully compare these four views, and in a way that is different from considerations put forward by McCall & Lowe (2003), Miller (2005a), and Hirsch (2008), we will see that some of these traditional enemies (namely, the perdurantist worm view and the various endurantist theories) are actually very much alike, and that some alleged points of substantive dispute fall prey to closer scrutiny.

A good way to see how the perdurantist worm view and its alleged opponents work is by examining how these theories handle the case of intrinsic change through time. My neighbour Cyrano, for instance, had a big nose, but after some time he discovered that there is a new easy, painless and very quick plastic surgery method that can replace his big nose with a small one, he then decided to undergo the procedure, and consequently he now has a small nose. In this case, Cyrano then undergoes intrinsic change—he first has a big nose and then a small one. What the worm view theorists claim here is that Cyrano is a *space-time worm*, that is, a

[1]See also Bennett (2008) and Chalmers (2008).

temporally extended entity that has temporal parts at every time at which it exists, and that his having of different incompatible properties at different times is a matter of him having different temporal parts at different times that have *simpliciter* the incompatible properties. Temporal parts are entities just like Cyrano, only temporally smaller, but not necessarily instantaneous—they can be temporally extended exactly as Cyrano is. Thus, according to the worm view, people are spatioztemporally extended worms that have temporal parts, and the phenomenon of qualitative intrinsic change over time is handled in terms of the having of qualitatively different temporal parts at different times.

Endurantism, on the other hand, claims that objects and people like Cyrano persist through time by being *wholly* present at all times at which they exist—they are thus *multiply located* at various times. Here is how one could start to try to understand this claim:



Such a picture of what endurantism is or could be is (would be) a strange one. Try to consider the analogous spatial picture: an object like a person 'multiply located' at several places in a conference room, say. Imagine an entire audience at your talk, only composed of one 'multiply located' person that would thus occupy the whole room. Since material objects are not universals, such a claim clearly does sound unacceptable, and the more natural thing to say would be that there is not one single object but a series of different objects laid before one's eyes. Since we are working here under an eternalist hypothesis, the endurantist picture about how Cyrano persists through time would then be as strange as in the analogous spatial case.

None of this shows that there is a problem with endurantism. Rather, it shows that the picture above and the way this picture suggests we should understand how endurantism works is a bad one. To understand why, and to better understand what the endurantist claim amounts to, let us see how endurantists typically answer an often-raised objection against their view: the Lewis-style objection from temporary intrinsics. Following endurantism, Cyrano at t_1 is numerically identical to Cyrano at t_6. At t_1, he has a big nose, at t_6, he has a small nose. But if we follow Leibniz Law, then if Cyrano at t_1 and Cyrano at t_6 are numerically identical then they should have all the same properties. But this leads to the untenable claim that Cyrano, the very same object existing at t_1 and t_6, has the two incompatible properties of having a big nose *and* having a small nose. David Lewis once considered this problem to be "the principal and decisive objection against endurance" (Lewis 1986, p. 203). To answer any worries about the having of incompatible properties, *perdurantists* defend a claim that is revisionary about what it is that has the incompatible properties: temporal parts, rather than 'whole' people—since the

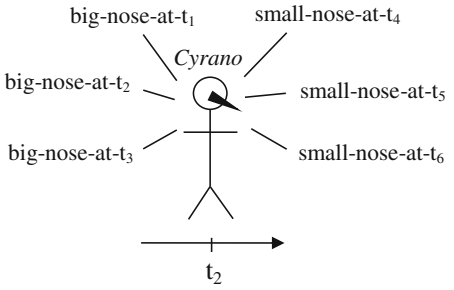
different temporal parts that compose a single space-time worm are not numerically identical, no threat of contradiction arises here. *Endurantists* typically appeal to at least two different strategies to answer the Lewisian worry. The first is Peter Van Inwagen (1985)'s strategy which is revisionary not about what it is that has the incompatible properties, but about the properties themselves. According to such a view, properties are always *time-indexed* and consequently Cyrano does not exemplify two incompatible properties such as "having a big nose" and "having a small nose", but rather he has the time-indexed properties "having-a-big-nose-at-t_1" and "having-a-small-nose-at-t_6" which are perfectly compatible. Contradiction avoided.

There is a follow-up to this argument that perdurantists often raise: granted, there is no problem in the having of the two time-indexed properties, but even if we grant that there are such properties, there still also are non-indexed properties like "having a big nose" and, if that's the case, the contradiction has not been avoided because even if Cyrano has at different times non-contradictory time-indexed properties, he also has the non-indexed properties—and so trouble comes back through the back door.

I find this perdurantist reaction somewhat strange. What it commits one to is to claim that Cyrano's having of a property is his having of it *simpliciter* without any disguised relations to times being involved. The reason why such a reaction is a strange one, coming from a perdurantist, is that while it is true that endurantism cannot accommodate this claim, the perdurantist (worm) view does not accommodate it either. Indeed, according to perdurantism Cyrano *also* has his properties only via a *temporalizing device* (think of the parallel with the *unifying device* we encountered in Chap. 1, §3): Cyrano, the temporally extended space-time worm, does not have a big nose. He can only be said to have this property by having a temporal part that has it. As a consequence, neither endurantism nor the perdurantist worm view can defend the claim that *Cyrano* has his temporary intrinsic properties *simpliciter*.[2] Perdurantists temporalize objects, while endurantists temporalize properties, and despite Lewis's objecting to the use of temporalized properties, and Van Inwagen's objecting to the use of temporalized objects (see for instance Van Inwagen (1985, p. 194)), what both views do is use a theoretical temporalizing device that plays the same theoretical role of making it possible for *Cyrano* to have properties. More precisely, the device "to be a t_n-part of" plays here the same overall theoretical role, and helps to solve the same problem, as the device "-at-t_n". As in the case of the bundle theory, the substratum theory, relationism, and substantivalism, we step here again on these theoretical tools I call "problem-solvers". A problem-solver is a primitive of a theory that is there to solve a problem. Both perdurantists and endurantists account for the phenomenon of intrinsic change through time by using their primitives: the temporalization of objects, or the temporalization of properties. At the same crucial places, both views introduce a tool with the same function: to avoid any contradiction arising from Cyrano's persisting

[2]It is true that only the perdurantist worm view allows for *something* (but *not* Cyrano) to have temporary intrinsic properties *simpliciter*, namely, temporal parts of Cyrano. I will come back to this later.

through time and having incompatible properties. Thus, both endurantism and perdurantism use a theoretical temporalizing device in order to avoid the threat of contradiction from the having of temporary intrinsic properties, and so, not only should endurantists be allowed to use *their* temporalizing device by their opponents, but also we have just made a first step towards the claim that the difference between endurantism and perdurantism is perhaps not as big as one would initially think. Furthermore, what we learn here is how we should picture endurantism correctly:



Following Peter Van Inwagen's way of drawing the picture, if t_2 is the present time, Cyrano is depicted as having a big nose, but he also has all of his time-indexed properties, which he has at all times at which he exists. This latter point is an important one, and we shall now see it brings us closer to the idea that endurantism and the perdurantist worm view are more similar than one could have thought. To better understand why, let us examine the traditional 'no-change objection' to the worm view.

The worm view's solution to the problem of the changing of intrinsic properties through time has raised a worry about its adequacy. Peter Simons for instance claims that the "four-dimensional [i.e. worm view] alternative is not an explanation of change but an elimination of it, since nothing survives the change which has the contrary properties" (Simons 2000a, p. 64). The problem here is that instead of accounting for one object's persistence and change through time, the perdurantist gives us a story about different objects (different temporal parts) that have different properties. Furthermore, if it is true that a temporal part of Cyrano has a big nose, it will always be true—such a fact cannot, accordingly to the worm view, ever change. One way to put this point as an objection is to charge perdurantism with the allegedly unpalatable task of defending a 'static' ontology where everything just seems to *be there* and where no object can ever *genuinely* change.

Now, the point of interest for us today is that this objection, if it were correct, would apply in exactly the same way to endurantism. Under endurantism as well as under perdurantism, the fact that Cyrano has the property of having-a-big-nose-at-t_1 is true at all times and can never change. All properties, according to indexicalist endurantism, are time-indexed, and consequently any property that Cyrano has, he has at all times at which he exists. Interestingly, he has at t_1 *the very same properties* that he has at t_5, and so, the friend of the 'no-change objection' can claim, he does *not* undergo genuine change between t_1 and t_5 (and so on). My aim here is of course not to object to endurantism. Like many others, I believe that these worries

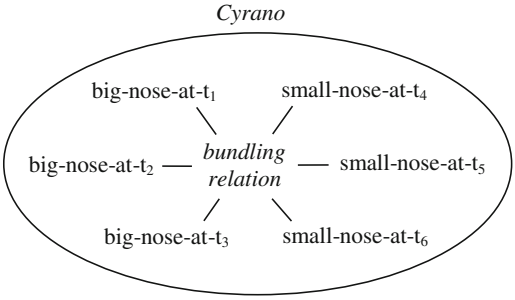
are easily answered. What is at the centre of my interest here is that *if* the 'no-change objection' applies, it applies equally to both endurantism and perdurantism (and *if* it does not apply, it does not apply to either of the two views).

§3. We have seen above the case of temporary intrinsics, which was supposed to be an objection to endurantism and a reason to favour the perdurantist worm view, but we have seen that it is not, and we have also seen the case of the no-change objection, which was supposed to be an objection to the perdurantist worm view and a reason to favour endurantism, but it is not—either both theories are guilty or neither is. (Actually, if anyone is guilty here, it is eternalism.) Thus, until now we have seen two steps towards the claim that the perdurantist worm view and endurantism work in a very similar way in some crucial places of alleged disagreement. Let us now see another traditional problem that is typically said to favour perdurantism over endurantism, and see the way the two views handle it: the Statue and the Lump case.

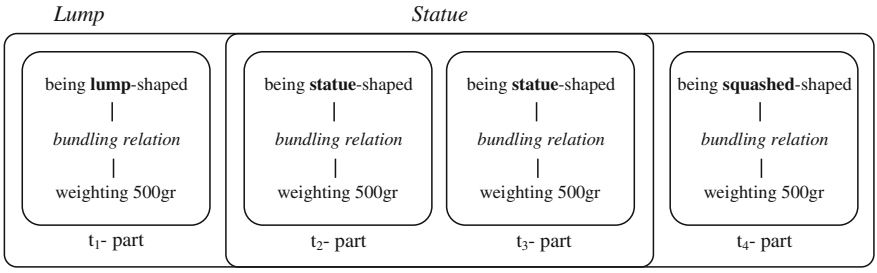
At t_1, there is a lump of clay that at t_2 an artist forms into a statue. A statue is thus created at t_2. Let us suppose that it persists until some later time, say t_3, and is then destroyed (squashed). Consequently, at some time after its destruction, at t_4, the statue does not exist anymore but the lump of clay still does: it persists from t_1 to t_4 where it existed at t_1 in some (let's say cubic) form, then it was shaped into the form of a statue and, after the destruction, it was shaped again into some other squashed form. The traditional puzzle consists in the fact that in the interval of time from t_2 to t_3, the lump of clay and the statue are one and the same object (they have the same form, the same location, they are made up of the same particles) but that if they were one and the same object, they should, following Leibniz Law, share all their properties, which is not the case since the lump of clay has, for instance, the historical property of being cubical at t_1 that the statue has not. So, after all, the statue and the lump of clay are different objects. But then, it seems that we have a situation where two distinct objects coincide between t_2 and t_3, which is typically supposed to be an unacceptable claim (as Lewis puts it: if the lump weights 500 g, and the statue weights 500 g, and if both objects are there between t_2 and t_3, why don't we have in this interval of time something that weights 1000 g?). Traditionally, perdurantists use this case to show that their view is superior to endurantism. Indeed, perdurantism has a simple reply: the t_2-part and the t_3-part of the statue *are* numerically identical, respectively, to the t_2-part and the t_3-part of the lump of clay. The t_2-part of the statue and the t_2-part of the lump of clay *do* share all of their properties, and relevantly, they don't have any different historical properties such as "being cubic at t_1" because none of them existed at t_1. But this does not entail that the *statue* and the *lump of clay* (the worms) are identical since for instance the lump of clay has parts at t_1 but the statue does not. So they are not identical but they share identical temporal parts: they temporally overlap. Consequently, following the perdurantist worm view, the case of 'coincident entities' is no more remarkable than the spatial case of two overlapping roads, one of them being a sub-segment of the other (see Sider 2001a, pp. 6 and 152). Endurantists, on the other hand, do not seem to be able to face this puzzle as easily, since it is the entire statue, and not a part of it, that is wholly present at t_2 or t_3, since the same holds for the lump of clay, and since they are distinct objects because they do not share all of their properties, the endurantist has to endorse the claim that, between t_2 and t_3,

there are two numerically distinct objects that coincide. This is why the case of the Statue and the Lump (as well as similar cases involving coincident entities) is typically taken to be a strong reason to favour the perdurantist view over endurantism.

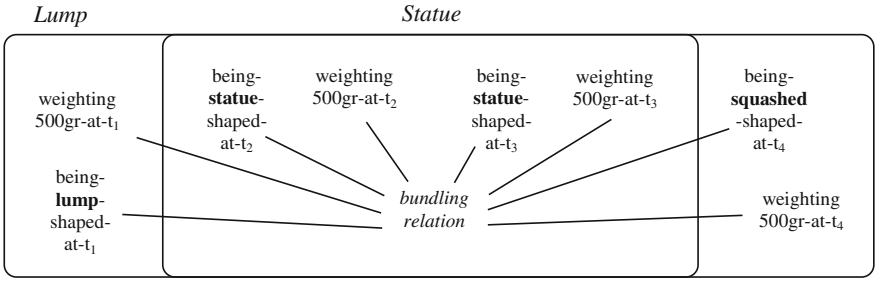
Before we see if this is really so, let us concentrate more carefully on how endurantism works and let us try to be more precise about the theory's structure. To be more precise, we need to stop drawing the endurantist picture in terms of drawings of people with big noses, and consider what the picture looks like when representing the fundamental components of the nature of Cyrano. To echo our discussion of Chap. 1, I shall consider here the two main options: either Cyrano is a bundle of properties, or he is a bare particular (substratum) that instantiates properties—with no surprise now, we shall see that choosing one rather the other will not make much of a difference. Under the view which is a combination of eternalism, endurantism, indexicalism, and the bundle theory, Cyrano is a bundle of properties (that is, *all* of his time-indexed properties) glued together by a special primitive bundling relation whose theoretical role is to bundle together properties in order to make particulars such as Cyrano.



Now, how can such a view handle the case of the Statue and the Lump? The *perdurantist* bundle-theoretic picture of the case is the following, where the bundle Statue is simply a *sub-bundle* of the bundle Lump—this is how, in terms of the bundle theory, we get the notion of temporal overlap used above by the perdurantist.



Having learned how the *endurantist* (indexicalist) picture should look like, we can now see how it can treat this case:



Lump is a bundle of time-indexed properties, Statue is a bundle of time-indexed properties, and one of the bundles is simply a sub-bundle of the other. Thus, such a picture provides a nice surprise for the endurantist: to account for this case, she can use the very same strategy that the perdurantist has been using all along. Exactly as under the perdurantist worm view, the bundle Statue is a sub-bundle of the bundle Lump, and consequently we get here an implementation of the notion of temporal overlap. Under both perdurantism and endurantism, this notion gives us the means to talk about two objects (if you want, you can say two "coincident" objects, but they are not coincident in any objectionable way, there are two objects in the perfectly acceptable sense in which there are two objects where there is a common part of two Siamese twins), but also to talk about one object (the common part of the two Siamese twins is one). Both views can thus equally well account for talk of two objects and talk of one object in a non-objectionable way. The endurantist can simply appeal to the same strategy the worm view does.

Nothing hinges here on the choice of the bundle theory, since the same treatment can be given under both perdurantism and endurantism if one embraces the substratum theory as well. According to the substratum (or 'bare particulars') theory, Cyrano is not only a bundle of properties, rather his properties inhere in a substratum that exemplifies them and unifies them in order to make a (thick) particular. With respect to my present concerns, this difference does not matter: whether it is a substratum that unifies the properties in order to make a particular, or whether they are united by the bundling relation, the resulting structure is such that it can easily accommodate the notion of temporal overlap as it is needed to provide a satisfactory treatment of the Statue and Lump case. This adds up to my discussion from Chap. 1: here again, we see that the bundle theory and the substratum theory can play the same role in the same way.

It took us a little time to get there, since we needed to be careful about clarifying how endurantism is to be understood, but here we are: first, endurantists can handle as easily as perdurantists the case of the Statue and Lump (as well as all similar cases involving so-called 'coincident entities'), and second, the general and more important truth is that the difference between the perdurantist worm view and endurantism is getting smaller and smaller. Indeed, contrary to the way these two alleged enemies are usually presented, *both views* implement the notion of *temporal overlap*.

§4. We have seen that endurantists can easily face some of the strongest objections that are often raised against their view, namely those that arise from

apparent cases of coincident entities such as the case of the Statue and the Lump. Generalizing, we can conclude that endurantism and the perdurantist worm view have the same explanatory power with respect to the puzzle cases involving coincidence, and this completes another important step towards the claim that the difference between these two views is much smaller than what is usually thought. To sum up, we have gone through four steps:

i. both views have to use a temporalizing device ("to be a t_n-part of" and "-at-t_n") in order to be able to say that Cyrano has a big nose or a small nose, and none of them can say that Cyrano has a big nose or a small nose *simpliciter*
ii. both views have to (and can) equally face the 'no-change objection'
iii. both views implement the notion of a temporal part (temporal overlap)
iv. by using the notion of a temporal part (temporal overlap) both views can equally well provide a satisfactory treatment of puzzle cases involving coincidence such as the Statue and Lump case

All of the four steps above were supposed to constitute the main differences between the two views, and all four have even been considered as being decisive in favour of one of the views over the other. (To provide only one reference for each step: Lewis (1986, p. 203) at one point thought that (i) was decisive against endurantism; Simons (2000a, p. 64) thinks that (ii) is decisive against perdurantism; Van Inwagen (1981, p. 90) thinks that (iii) is decisive against perdurantism since the notion of a temporal part is unintelligible; Sider (2001a, Chap. 5) thinks that (iv) is decisive against endurantism.) So, if these four central points of alleged disagreement between the two views collapse, doesn't it in the end turn out that there is *no* difference *at all* between the two views, and that they only are some sort of terminological variants of each other? No. Such a conclusion cannot be drawn from the considerations I put forward in this chapter, and I believe that it is also an incorrect one, because there are some genuine and substantive differences between the two theories.

A first and important point of departure between endurantism and the perdurantist worm view is that while it is true that neither of them can say that Cyrano has a big nose or a small nose *simpliciter*, the worm view *can* say that *something* has a big nose or a small nose *simpliciter* (i.e. one of his temporal parts). A second difference between the two competitors is that they are structurally different: this is easily seen if one uses the substratum theory, for the perdurantist worm view will claim that there is one substratum per time that unifies the properties had by Cyrano at that time, while the endurantist view will claim that there is only one substratum that unifies all of the properties that Cyrano ever has (a parallel claim can of course be made under the bundle theory as well, appealing to one or to several relations of compresence, as we have seen in Chap. 1). This justifies the *endurantist* claim that material objects persist through time by being *numerically identical at different times*, while this is how perdurantists account for the claim that *nothing is ever numerically identical at different times* and that objects persist through time by having temporal parts. There is a link between these two differences between our two theories, since it is only because of their different structure that they exhibit a difference in the way the two views can or cannot claim that *something* has

properties such as having a big nose *simpliciter*. Thus, what we have learned is *not* that the perdurantist worm view and endurantism are somehow, on a general level, equivalent; rather, we have seen that some traditionally important points of departure actually show how similar the two views are, but that they also are different with respect to some other points. It would thus be incorrect to say that they are 'equivalent' or 'merely terminological variants' in general, while it *is* correct to say that this is true to some (important!) extent.

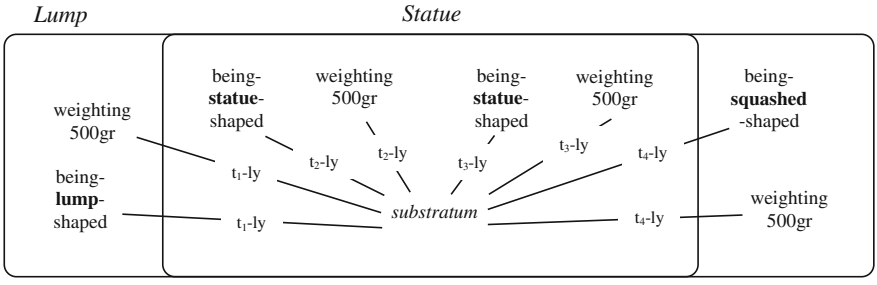
§5. In the discussion above, I have used the indexicalist version of endurantism, but this is not the only strategy endurantists can appeal to in order to answer the problem from temporary intrinsics. Importantly, there is the *adverbialist* solution according to which one should not temporally modify the properties Cyrano has, but the having of these properties. Under adverbialism, "Cyrano has a big nose at t_1" is to be analyzed as "Cyrano has-at-t_1 a big nose" or as "Cyrano has t_1-ly a big nose" as Johnston (1987) more elegantly puts it. In this view, there is not just the having of a property, there is always t-ly having (or having-at-t) of a property. Any worries about the having of temporary intrinsic incompatible properties are thus easily dissolved, since while it is true that Cyrano has a big nose at t_1 and has a small nose at t_4, and so he has both the incompatible properties, he has the former t_1-ly and the latter t_4-ly and this is how contradiction is avoided. (I focus here on Johnston's brand of adverbialism because it seems to me to be the best—and the most straightforward—version of the view. Other versions can be found in Lowe (1987, 1988) and Haslanger (2003). I critically discuss these versions in Benovsky (2006, Part I, Chap. 4, §16–21).

With respect to my discussion above, there is one important difference between adverbialist endurantism and indexicalist endurantism: only indexicalism, but not adverbialism, is compatible with the bundle theory. The *substratum* theorist, if she wants to be an adverbialist, can say that there are three components in her view: a substratum, its properties, and a relation of exemplification that holds between the substratum and the properties (and which is time-indexed, as the adverbialist view requires it). The *bundle* theorist, on the other hand, does not have room for such a picture in her ontology, since she does not postulate a substratum that needs to be related by a special relation to its properties—rather, in her view, such intermediaries should be avoided and so she cannot be an adverbialist since there simply is no suitable place to put the adverbialist index.[3] This, of course, holds only for a very special version of the substratum theory, namely an unpopular version of this view which insists on there really being this third component in the theory: the (time-indexed) relation of exemplification between the substratum and the properties it has. Many substratum theorists often rightly agree that this is a bad version of their view, among other reasons because of Bradley-like regresses and related worries. As Sider (2006) in his recent defence of substrata insists, this relation of exemplification should not be taken too seriously, in the sense in which it is often claimed that exemplification is not a genuine relation, that it is a "non-relational

[3]If one were to put the index on the bundling relation, it would straightforwardly become a *perdurantist* view.

tie", and that we shouldn't "reify" exemplification (see, for instance, Lewis 1983, pp. 351–355). To my mind, these worries are justified, and relevant to my discussion here, if the friend of the substratum theory follows these recommendations, she then cannot be an adverbialist for the simple reason that if she takes away from her view the ontologically significant relation of exemplification there will be no good place to put the adverbialist index anymore. Only if she is not impressed by the troubles that arise when one takes exemplification ontologically seriously as a relation (that one can put an index on), does she have the option of holding an endurantist-adverbialist-substratist view. (To my mind, this makes adverbialism an unpalatable solution to the problem of persistence through time in the first place.)

This being said, let us now see how adverbialism compares to indexicalism and to the perdurantist worm view. The first point of similarity between these views holds: exactly as it was the case for endurantist indexicalism and for the perdurantist worm view, adverbialism also has to use a temporalizing device ("t_n-ly") in order to be able to say that Cyrano has a big nose or a small nose: all three views thus can*not* say that Cyrano has a big nose or a small nose *simpliciter*. Furthermore, since the adverbialist theory is here combined with eternalism it also has to (and easily can) face the 'no-change objection' for the very same reasons we have already given in the case of indexicalism (and the perdurantist worm view). Interestingly, adverbialism also implements the notion of temporal overlap and, exactly like the two other views, it can equally well provide a satisfactory treatment of cases such as the Statue and Lump case, as the following figure shows—analogously to what we have seen in the indexicalist's case.

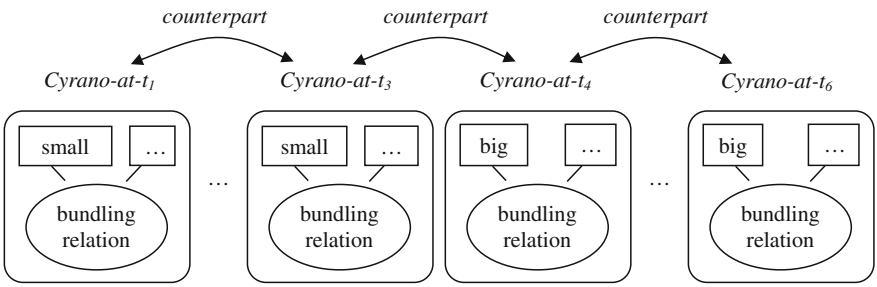


As a consequence, we can affirm that endurantism-adverbialism-eternalism-substratism is *not* very different from the perdurantist worm view and the endurantist indexicalist view with respect to the same (important) points of alleged disagreement between endurantism and perdurantism, while it *does* differ from the perdurantist worm view for the same two reasons we have seen above concerning the difference between indexicalism and the worm view—the additional difference being here that only the worm view, but not adverbialism, is compatible with the bundle theory.

§6. Finally, I now turn to an eternalist theory of persistence through time that *is* different from the other three views with respect to the four steps examined above:

the perdurantist *stage view*. I shall first go through the four steps (i)–(iv), and then ask what kind of difference we are dealing with here: metaphysical, or purely semantic/linguistic?

The stage view's account of persistence and change over time is the following (I use the bundle theory, but like in the case of the worm view and of endurantist indexicalism and *un*like in the case of adverbialism, using the substratum theory would make no relevant difference here—strengthening again my claim from Chap. 1):



According to this view, a person like Cyrano exists only at *one* time and is an instantaneous entity (an instantaneous 'stage'), and it persists through time by having different temporal counterparts at other times. The ordinary object we refer to as Cyrano is not a four-dimensional (temporally extended) entity, rather, there is a series of stages interconnected by a counterpart relation, and ordinary objects are conceived of as being the stages rather than the whole composed of them. Strictly speaking, the different Cyranos are only momentary entities, but they are nevertheless said to persist through time by having counterparts at other times. However, the stage view is still a four-dimensionalist view since it does not deny the existence of temporally extended objects—the four-dimensional entities that are aggregates of stages—they exist as well as the stages do. It's just that, according to the stage view, the objects we ordinarily name and quantify over are stages rather than worms.

Let us now see how the stage view behaves with respect to the four steps (i)–(iv) above.

The first point of departure from the three other views is already easily seen when it comes to the having of temporary intrinsic properties *simpliciter*. The stage view, unlike the perdurantist worm view and the two endurantist views, can guarantee the having of temporary intrinsic properties *simpliciter* by ordinary objects themselves. This is easily achieved since, according to this view, ordinary objects like Cyrano(-at-t_1) are (instantaneous) stages, and those can have properties *simpliciter* without making them to be relations to times. So such a view allows Cyrano to have *simpliciter* the property of having a big nose (in a non-derivative way, contrary to what the worm view has to say).

What about the second step, the 'no-change objection'? It seems that here also, the stage view behaves differently than the three other views—the fact that it appeals to different counterpart-related objects to provide an account of persistence

makes it *weaker* against the objection. Let us remember shortly how it goes: perdurantism (worm view) does not account for genuine change in persisting material objects like Cyrano, because it tells us a story about different objects (his different temporal parts) existing at different times and having different properties, but none of those objects can change. And similarly for endurantism (take the indexicalist version here): the fact that Cyrano has the property of having-a-big-nose-at-t_1 is eternally (always) true and can never change; indeed, any property that Cyrano has, he has it always and forever, and so he has at a time t_1 the very same properties that he has at a later time t_2, and so the 'no-change objection' goes, he does not undergo genuine change.

As I already mentioned above, and as many have argued, I don't think that either of these views really has anything to fear here, since all have an equally adequate reply to the objection. In short, here it is: the perdurantist can say that there is something that changes, namely the four-dimensional Cyrano who is composed of all of his temporal parts. Once one of his parts has any intrinsic property, it cannot change, and it will always be true that it (tenselessly) has this property, but the four-dimensional entity can be said to undergo a change by having different parts at different times. Change is simply the having of different properties at different times, and the perdurantist's worm can easily accommodate this claim. And so can (obviously) the endurantist.

But if one endorses the stage view, such a reply seems unavailable—for there is no *one* thing that ever has the different properties. The worm view theorist claims that the temporally extended Cyrano has them in a derivative way, and the endurantist claims that he has different time-indexed properties, or that he has them t_n-ly, but the defender of the stage view does not have room for any of this in her theory: she cannot show anything that could be said to undergo a change, even in a derivative way. Of course, she can say that a certain stage, say at t_1, is F and will be ¬F at t_2 in virtue of being a temporal counterpart of another stage existing at t_2 that is ¬F. This could maybe sound like a solution, but it is not: for these two stages are just two completely different things. As Mellor (1998, p. 89) puts it, "change needs identity as well as difference". But there is only difference in the stage view, there are only different things with different properties.

This is not the end of the story, of course. For what is involved here, and what the stage view theorist's reply will criticize, is a version of the 'Humphrey objection' applied to temporal counterpart theory. The objection runs as follows: if Cyrano says now that he will visit Roxanne tomorrow, then the sentence turns out to be true iff he'll visit Roxanne tomorrow. But this is, according to the stage view, simply impossible, because the person who says now that he'll visit Roxanne tomorrow is a stage, a momentary entity that will not itself persist until tomorrow and thus, will not be able to visit anyone. Cyrano, the person who is doing the speaking, is simply not identical, in any sense, to the person who's supposed to do the visiting. Granted, Cyrano has a counterpart tomorrow that'll visit (or not) Roxanne. But whatever the counterpart relation is, it is not identity. So, the objection goes, if Cyrano says he'll visit Roxanne tomorrow, why would Roxanne care that someone else, similar to Cyrano and linked to him by a counterpart

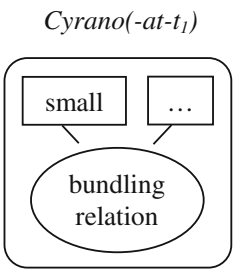
relation, will visit her? Note that the "someone else" claim is very strong here: if one endorses the stage view, there is no sense in which Cyrano from today is identical to Cyrano tomorrow. If one generalizes this objection, one can simply claim that it denies persistence altogether. For the stage view ontology provides us only with instantaneous entities, and aggregates of those, but it rejects 'worms' in the sense that it rejects the view that ordinary objects like people are four-dimensional entities. The stage view's persistence, the objector says, is not genuine persistence.

Sider defends the stage view against this objection as follows: "[It] is wrong to say that the stage view denies that 'You will do it' means that you will do it. 'Ted was once a boy' attributes a certain temporal property, the property of once being a boy, to me, not to anyone else. Of course, the stage view does analyse my having this property as involving the boyhood of another object, but I am the one with the temporal property, which is the important thing. The stage view is consistent with stages having temporal properties; it's just that temporal properties are given a counterpart theoretic analysis" (Sider 2001a, p. 195).

But this reply is not likely to give satisfaction to the objector. Granted, the stage view is consistent with stages having temporal properties, but they are not the ones we want. To take Sider's example, if we say "Ted was once a boy", we are ascribing a certain temporal property to Ted (who exists now). But if we want to endorse the stage view, it is not the property of "once being a boy", but rather, the property of "once there being a counterpart of Ted that is a boy". If the stage view theorist allows these two properties to be equivalent then she is mistaken—for if it is the former that we ascribe to Ted, we are speaking solely about Ted, but if we ascribe him the latter, we are speaking about Ted and someone else, and that makes all the difference. Sider's response can only appear to be satisfactory if one takes the expression "once being a boy" to be a suitable paraphrase of the expression "once there being a counterpart of x that is a boy", but such a strategy, objectors like Sally Haslanger will claim "strains the limits of credibility" (Haslanger 2003, p. 337).

Although I have sympathies with the objector, my point here is not to claim that we should reject the stage view because of this objection. Rather I only wish to claim that the stage view has more to do than the worm view or any of the versions of endurantism in order to answer the no-change objection (step (ii)), and that its reply has to be different, since it cannot appeal to any *one* object having different properties at different times, in the way the two other views do it.

With respect to step (i), the stage view has an advantage over the other competing views, and with respect to step (ii), it seems, on the contrary, to be in a weaker position. This should not be very surprising, since the stage view is also structurally different from its competitors (this is step (iii)). As we have seen, according to the stage view, a person like Cyrano is 'no more' than this:



which is quite different from the worm view and both versions of endurantism. To put it simply, the three other views claim that Cyrano is 'bigger'—he is a bundle not only of properties he has at one time, but of *all* of his properties he ever has. And this salient difference in structure will also make the stage view behave very differently with respect to step (iv): the case of the Statue and the Lump. While the crucial notion that is appealed to by the worm view and endurantism in their treatment of the puzzle is the notion of overlap, nothing similar is either available or needed if one embraces the stage view, since there is nothing temporally extended that could be said to overlap. At a time t_2, for instance, there is only one instantaneous entity that is a statue made out of a lump of clay but there are not two coincident objects at this time, since the reason for thinking that there could be two different objects was that they were suspected to have distinct historical properties like "being cube-shaped in the past" or "having existed at t_1", but no instantaneous entity has any such properties. It can be said to have them by having different temporal counterparts at different times, but the counterpart relation being flexible (context dependent), it will be able to have different counterparts qua Statue than it has qua Lump—so what we have is just one object that has different counterparts under different counterpart relations and there is no threat of ending up with coincident entities.

To sum up: the stage view is different from the three other views with respect to all four steps (i)–(iv). It behaves better in the case of temporary intrinsic properties, it is weaker with respect to the no-change objection, it is structurally different, and it provides a different treatment of the Statue and Lump case.

The diagnosis of why exactly this view is different from the others is readily at hand: it takes objects like Cyrano to exist at only one single time, while all the competing views take them to exist at more than one time. To put it in a more objection-like way: it actually denies genuine persistence through time instead of providing an account of it.

One way to support this objection is to insist on the importance of the fact that the stage view is weaker with respect to the no-change objection, as we have seen above, and to insist that the 'Humphrey objection' to temporal counterpart theory succeeds. But there is also another (but related) way to see this defect of the stage view: the view does not allow ordinary objects to do the things they typically can do. People, like Cyrano, are stages. But stages are instantaneous entities, they do not have temporal extent. The unwelcome consequence of this is that people cannot

do many of the things we would expect them to be able to do. For instance, it seems that a person should normally be able to utter a sentence. But, on the stage view, this turns out to be impossible, strictly speaking: the utterance of a sentence takes some time and a stage does not last long enough to make such a performance. Or, normally, Cyrano can run, but again, not according to the stage view; strictly speaking nobody can run because a person is an instantaneous entity and running takes time.

The obvious reply of the stage view theorist to this is that Cyrano can utter a sentence and run because he has counterparts at 'neighbouring' times and if we take several counterparts together, they can achieve such a performance. It takes more than one single stage to speak or to run.

But then, what do we really mean when we say that a person runs? What do we refer to by "this person"? We have seen that it seems that if we refer to the instantaneous stage (as we should, if we follow Sider's recommendations: space-time worms "are not ordinarily named or quantified over" (Sider 2001a, p. 191)) it is impossible for our sentence to be true (an instantaneous entity does not have enough time to run).

Maybe we refer to a sum of successive person-counterparts, which is a thing that lasts long enough to do the performance. But what is this sum? I see two possibilities: first, that it is a set of numerically distinct entities (the distinct temporal counterparts), or second, that it is a whole composed of the different counterparts. The first possibility seems really unpalatable: the view according to which a set of distinct objects can run would be hard to defend. We are then left with the second possibility; but this just amounts to embracing the worm view, for the thing that has the properties we are interested in (running, speaking, and so on) is a temporally extended four-dimensional entity—and so, those properties are really had by a 'worm' rather than by a stage. So, since this way out (a way that nobody takes, as far as I know) is closed, the stage theorist will have no choice but stick to his original claim: "this person" refers to an instantaneous stage, and it has the property of running in virtue of having temporal counterparts at neighbour times—and this is how a person can run. Very well. But what this claim commits the stage view theorist to, is to endorse the further claim that since the person has the property of running in virtue of its relations to other persons (his past and future counterparts), this property turns out to be extrinsic, contrary to what we'd usually say. And for the same reasons, a lot of properties that we usually take to be intrinsic turn out to be extrinsic, according to the stage view. So if one wants to account for the fact that people can speak and run, and that the properties involved here are intrinsic, one should embrace the worm view or endurantism rather than the stage view.

Again, this is not the end of the story. For the stage view theorist will defend his view here by simply biting the bullet and accepting that most properties we thought to be intrinsic are, in fact, extrinsic.[4] But, again, my point is not here to establish

[4]Both Ted Sider and Achille Varzi do accept this consequence of the stage view (personal communications, 2005); see also Sider (2001c).

whether we should accept the stage view or not, although I have been offering reasons that point towards rejecting it. The point I wish to make by raising the objections above is to show in what respects the stage view is different from its competitors, and I think that this has been established.

Still, one might wonder what kind of difference this really is. Is it a metaphysical difference? It certainly is not a difference in what there is, since all the views are eternalist and postulate the existence of the same distribution of matter across space-time, and the stage view does not deny the existence of mereological sums of stages that correspond to the worm view's space-time worms.[5] So the difference is not one in what there is (contrary to what is the case in the presentism-eternalism controversy, which will be the concern of Chap. 3), but in the analysis of what ordinary objects like tables or people are. Is this a metaphysical or a semantic difference? In a sense, as Sider himself claims, it seems only to be a semantic/linguistic one, since the disagreement only seems to be about ordinary language terms and reference—a disagreement located in what we usually name and quantify over when we make claims about ordinary objects.

But, as Parsons (2004, p. 3) points out, rightly I think, metaphysical questions are not only questions about *what* there is, but also about *how* things are. To take his example: "Does time pass?" is as much a metaphysical question as "Does the future exist?" It is true, Parsons says, that the worm view and the stage view agree on the stuff there is, but it doesn't follow that they agree on all metaphysical questions—like the question of what the nature of tables and people is, that is, *how* they are. The question whether, for example, I am three or four-dimensional *is* a metaphysical one. Or, the two views do not provide the same answer to the question: "how many people are there crossing the street when Cyrano crosses a street?"—indeed, there are *much* more people there according to the stage view (as many as there are instants, or infinitely many if time is continuous) than according to what the other views say. So it seems that the disagreement between the stage view and the other competing views I have discussed is not merely semantic/linguistic but genuinely metaphysical, and that it is about whether ordinary objects are best conceived of as time-bound (momentary) or extended in time.

§7. We have seen that the debate between endurantists and perdurantists is, to a large extent, verbal and that there is much less substantive disagreement than we could have thought. But, importantly, genuine differences and room for substantive disputes remain. I would like to suggest that this is quite representative of the state of metaphysics, given the ongoing meta-metaphysical debate: some areas of metaphysics, that we thought were well explored and that we thought gave rise to competing incompatible views, turn out to be places of merely verbal disputes. But not all. And more: even 'inside' one particular debate, like the persistence one, there are merely verbal points and substantive ones. This is why I would like to emphasize something that is probably (hopefully) not very original: that we should

[5]Besides, it is likely that whatever the temporal counterpart relation is, it will turn out to be the same as the 'glue' relation that unifies the temporal parts of a single space-time worm.

not make any very general claims about the status of metaphysical debates, and not even about a status of one metaphysical debate, in order to claim that it is verbal or substantive or otherwise; rather, we should do first-level metaphysics in detail, examine the nature of particular detailed points of disagreement, and only then raise any meta-theoretical claims, like claims of metaphysical equivalence. In Chaps. 3 and 4, I will have more to say about this.

Chapter 3
Non-equivalent Metaphysical Theories

Abstract In Chap.1, I have discussed cases of equivalence between metaphysical theories, and in Chap.2, I have given an example and discussed the idea that, in some relevant cases, such an equivalence can be partial. In this chapter, I want to have a look at a case of two metaphysical theories—presentism and eternalism—which have been said to be equivalent, but which are not, and I will focus on the reasons why one might think they are, and on why these reasons are, in general, inadequate.

§1. In Chap. 1, I have discussed cases of equivalence between metaphysical theories, and in Chap. 2, I have given an example and discussed the idea that, in some relevant cases, such an equivalence can be partial. In this chapter, I want to have a look at a case of two metaphysical theories which have been said to be equivalent, but which are not, and I will focus on the reasons why one might thing they are, and on why these reasons are, in general, inadequate. The two theories I have in mind here are eternalism and presentism, and the general methodology for equivalence claims I will criticize is one close to Hirsch's "alternative languages" strategy (even though, of course, Hirsch himself did not apply this strategy to the debate between eternalists and presentists).

§2. Formulating eternalism is somewhat easier than formulating presentism, because as we shall see the problem of an alleged equivalence between the two views can already be apparent in some of the ways presentism is or can be formulated. Let us start with eternalism.

Eternalism is the doctrine about time which takes all times to exist and to be ontologically on a par—there is no ontological difference between past, present and future times. As Sider (2001a, p. 11) puts it: "Just as distant places are no less real for being spatially distant, distant times are no less real for being temporally distant". Thus, past and future objects exist, just as present objects do. In the eternalist's manner of speaking, future objects "exist", as well as present objects exist, in an atemporal sense of the verb; it is as if one were viewing the universe from God's standpoint and could contemplate all that happened, happens, and will happen laid before his eyes (Arthur Prior calls this "the tapestry view of time" (Prior 1996, p. 47)). Furthermore, eternalism is typically coupled with an additional thesis: not only the present time and presently existing objects have no ontological privilege, but also "now" is an indexical term just like "here".

J. Benovsky, *Meta-metaphysics*, Synthese Library 374,
DOI 10.1007/978-3-319-25334-3_3

Presentism, on the other hand, claims that only presently existing objects are real—"to exist", then, amounts to "to exist now". Typically, presentists are also 'serious tensers' drawing an important distinction between saying that past objects once existed and future objects will exist but only current objects exist. One could think, at a first glance, that presentism, thus formulated, is a non-starter (and this is where the meta-metaphysician can start to raise her eyebrows)—for how is one to understand the presentist's central claim, "The only things that exist are those that exist at present"? It seems there are two possibilities: either the first occurrence of "exist" in this claim is tensed or it is not. If it is, then it seems that presentism is an uninteresting truth ("The only things that exist now are those that exist at present"), and if it is not—that is, if "exist" is to be taken as a tenseless form of the verb meaning something like "existed, exist, or will exist"—then presentism seems to be obviously false. I shall come back to this shortly. But let us try to formulate presentism in such a way that it does not sound obviously true or obviously false: what this view wants to say is that there are fewer objects than those acknowledged by the eternalist. Thus, presentism is a thesis about what there is as, for instance, Zimmerman (1998, p. 210) puts it: "[T]here is only one largest class of all real things, and this class contains nothing that lies in the past or future. Presentism is, in fact, a thesis about the range of things to which one should be ontologically committed". Presentism is the view that only present objects exist. If we were to make an accurate list of all the things that exist, there wouldn't be a single non-present object on the list.

§3. At a first glance, and even with such rough formulations at hand, the two views could hardly look more dissimilar. But, some have claimed that, when one looks more closely, eternalism and presentism are equivalent. One can find various (and different) statements of this kind of equivalence claim in Lombard (1999), Callender (2000), and Dorato (2006)—as well as in many conversations during metaphysics workshops. I will not focus on the details of these various statements of this equivalence claim, I will rather be interested in the methodology. We have already encountered above the general reason one might have to think that the disagreement between presentists and eternalists is merely verbal: if "exists" means "exists now", then both sides agree that—trivially—everything that exists is present, and if "exists" means "existed, exist, or will exist", then both sides agree that dinosaurs exist. In both cases, both sides have the same entities on their inventory of what there is. Thus, one can claim that there is no substantive disagreement between the two sides after all.

More precisely, the idea is this: whatever one side claims, the other side can claim as well, provided they use their own meaning of "exist". Thus, if an eternalist says "Dinosaurs exist." where she uses "exist" in an atemporal sense, a presentist can say, using her own tensed talk, "There were, are, or will be dinosaurs." Both sides can there agree on the same truth, in their own ways of formulating it, and talk about dinosaurs. Where then, the critic asks, lies the alleged disagreement?

§4. I have to confess that I am less interested in presentism and eternalism here, and more interested in the methodology that lies behind such equivalence claims. I shall come back to presentism and eternalism below, but for now I would like to render more explicit the idea that those who claim that there is an equivalence between presentism and eternalism, for the kind of reasons given above, appeal to the idea that whatever one

side can say in its language the other side can claim as well in its own, different language. This is close to Hirsch's (2005, 2008) strategy (while Hirsch applies it to endurantism and perdurantism, and not to presentism and eternalism), which has, as a result, that many metaphysical debates turn out to be equivalent.

Hirsch says that a dispute is merely verbal if the controversial statements of alleged disagreement ("There are dinosaurs.") are most plausibly interpreted, using a principle of interpretative charity, as having different truth-conditions in different languages, so that each side of the debate is correct in its own language. Each side is here capable of expressing the same thing, but in its own language. Again, Hirsch does not apply this strategy to eternalism and presentism, but it's exactly what friends of such an equivalence claim are after. Sider (2001a, pp. 15–17), who does not himself defend such a claim, puts it very clearly like this: "The alleged disagreement is over quantified sentences such as 'there exists a dinosaur'. But the dispute would disappear if the presentist and eternalist meant different things by the quantifier. Suppose, for example, that what the eternalist means by 'there exists (atemporally) an x such that…' is what the presentist would express by a disjunction of combinations of tense operators and present tense quantifiers: 'WAS(∃x…) v∃x… v WILL(∃x…)'. Then, it might be claimed, the disagreement vanishes, for the presentist will accept the first disjunct of 'Either there was a dinosaur, or there is a dinosaur, or there will be a dinosaur'."

The idea is that each side has its own language—its own meaning of the quantifier—and so each side can express, in a different way, that there exists a dinosaur. Since, according to this strategy, one has to be charitable when interpreting both languages, both sides are right, in their own ways. For Hirsch, to put it simply, the only remaining question is then to decide which of the two languages is closer to ordinary English—it is this proximity that will decide which is better.

§5. The problem with such a strategy is that we get equivalence too cheaply. To put it bluntly, if two different theories say *prima facie* different things but if we can always translate or map what one says into the language of the other, we are systematically led to conclude that they are equivalent. But this is a methodological mistake. The fact that we can have a mapping or a translation of what one can say in one language to what one can say in the other language does not mean that the two theories are equivalent—it only means that they are able to express the same things (granted, this counts for *something*, but it's not equivalence).

Compare this strategy to what I have been doing in Chaps. 1 and 2, *à propos* of the bundle theory, the substratum theory, substantivalism, relationism, perdurantism, and endurantism. Even if we grant Hirsch and others the fact that there is a type of systematic translatability or mapping of one side's language into the other side's, when we examine in detail how the theories *work*, *what* they do, and *how* they do it, we realize that in some places they indeed appear to be equivalent (they do the same job in the same way), but in other places they do not. As we have seen in Chap. 2, even if there is a partial equivalence between endurantism and perdurantism, there remain some significant differences, and the fact that what one side can say what the other can say just as well in its own language is just not the most relevant fact here. We have examined in detail how the two competing theories

work, what they can do and what they cannot do, and what their explanatory power is—that's where we were able to identify places of equivalence, as well as places of genuine disagreement. The method of alternative languages, on the other hand, does not allow us to spot these differences, because it is too general, and because it does not pay enough attention to the inner workings of each side of the debate.

One thing Hirsch (2008) agrees on is that one should not make claims about metaphysics in general, saying, for instance, that, for principled reasons, all metaphysical debates are merely verbal. Instead, one has to limit such claims to particular debates. For Hirsch, as a consequence of this methodology, and of the strategy of alternative languages and interpretative charity, the endurantism versus perdurantism debate is merely verbal, but the platonism versus nominalism debate is not. But, while I, of course, agree that any such general claim about the entire metaphysical enterprise is very likely to be inadequate, I think that the method of alternative languages is still too general as well. The idea is this: we need to do first-level metaphysics, in order to be able to do meta-metaphysics. When asking whether metaphysical questions are substantive or merely verbal, one should not make any general alternative languages claims, rather one should carefully examine metaphysical questions and debates one by one *by doing first-order metaphysics* (I tried to provide some examples in Chaps. 1 and 2), and make decisions only case by case. Some ontological questions and debates are merely verbal, but some aren't, and as we have seen in Chap. 2, some debates can be partially merely verbal and partially substantive.

Thus, I do have sympathies with Hirsch's (2005, 2007, 2008) 'moderate anti-realist' view that claims that many metaphysical debates are merely verbal disputes where the disputants seem to claim different things but in fact they are making the same claims only formulated in different ways, or different languages. But I disagree about the way we can get to such a metaontological claim, and about when it applies correctly. Indeed, as I already insisted above, I think that when asking whether metaphysical disputes are merely verbal or whether they are genuine, the answer is "It depends". It depends on how the metaphysical theories at hand do their job—and *not* on their inter-translatability. My view is then different but close to Hirsch's and it is a 'moderate' one as well: on the one hand ontological disputes such as those we have seen above do not, sometimes, run very deep—but, importantly, there still is room for (sometimes only partial) genuine disagreement in some cases.

§6. Let us come back to presentism and eternalism. According to the equivalence claim, both views are able to express the same truths in their respective languages. These truths are truths about what there is, like "There are dinosaurs." But it is not clear that this claim always succeeds and that the two theories really do have the means to say the same things. Sider (2001a, pp. 15–16) identifies a possible case in which "[…] claims the eternalist *accepts* [map] to claims the presentist *rejects*. […] An eternalist who believes in sets would accept the claim that there exists a set containing a dinosaur and a computer, but the presentist will reject the disjunction: WAS(∃x, x is a set containing a dinosaur and a computer) or (∃x, x is a set containing a dinosaur and a computer) or WILL(∃x, x is a set containing a dinosaur and a computer). One can think informally of WAS(φ) as saying that at some past

time, t, φ is true at t. […] The first disjunct, then, says (informally) that at some time in the past there existed a set containing a dinosaur and a computer; the second says that there exists such a set at the present time, and the final disjunct says that at some future time, some such set exists. Since at no one time did there exist both a dinosaur and a computer, it follows that at no time will there exist a set containing a dinosaur and a computer (assuming that a set exists only if its members do). Thus, from a presentist's perspective, each of the three disjuncts is false." Of course, one can deny here the existence of sets (or fusions, or similar), but this is not what is at stake here: what this case shows is that, at least as a possibility, eternalism and presentism behave differently.

The alternative languages claim, as applied to eternalism and presentism, can then be weakened. But suppose, for the sake of argument, that we grant it. Even then, what this shows, at best, is that the two views have the same expressive power with respect to what there is. But this is not even the central issue in the eternalism versus presentism controversy. The central issue is the nature of time: eternalists say that time is space-like and that different times are like different places, while presentists deny that. Nothing in the translatability claim erases or minimizes this fundamental difference between the two competing views. This alone shows that the two views are not equivalent.

Compare this to the debate about the metaphysics of possible worlds. In this context, eternalism is akin to David Lewis' modal realism (i.e. all times are ontologically on a par; all worlds are ontologically on a par) and presentism is similar to some kind of actualism (i.e. only the present time is real; only the actual world is real). If, then, the argument for an equivalence between eternalism and presentism were to succeed, it would have to succeed in the case of the debate between modal realism and (some versions of) actualism. But however one spells out the difference between these two theories, it is clear that they *are* different. Modal realism claims that merely possible worlds are of the same ontological kind as the actual world is, and actualism denies that—exactly as eternalism claims that past and future times are of the same ontological kind as the present time is, and presentism denies this. Granted, there can be similarities between the two sides of the debate, and perhaps (but, we have seen at the beginning of this section that perhaps not) they, indeed, are able to express the same truths in their own respective languages. But none of this means that they are equivalent—they take possible worlds to be different things, or times to be different things.

§7. All of the above being said, let us see the best possible reason for rejecting the alleged equivalence claim between presentism and eternalism: an example where they are *not* equivalent following the methodology used in Chaps. 1 and 2, that is, by seeing that they do not work in the same way and that they do not have the same explanatory power. This will connect to the discussion of Chap. 2, since what I have in mind here is that eternalism, but not presentism, is compatible with the perdurantist worm view. If this is the case, it suffices to show that the two theories are different (and whether this is an advantage of eternalism or of presentism is a question we can leave open here). So, let us do some meta-metaphysics by doing first level metaphysics.

The perdurantist worm view's central claim is that ordinary material objects are aggregates of *temporal parts* and that they persist (perdure) from one time to another by having different temporal parts at different times. These temporal parts are 'time-bound', that is, they don't exist at more than one time. Aggregates of temporal parts (ordinary material objects, people,…) of course can and do exist at more than one time, but only by having numerically distinct temporal parts at different times. *Prima facie*, this type of perdurantism can be combined with both eternalism and presentism. The combination of perdurantism and eternalism is traditionally called "four-dimensionalism", and the combination of perdurantism and presentism can simply be called "presentist perdurantism". In what follows I will explore some aspects of this latter combination, while comparing it to the four-dimensionalist view, and I will argue that it falls prey to serious difficulties—again, only to show here that eternalism and presentism are different.

§8. Let us consider the presentist version of perdurantism. Such a view claims that an object that exists at the present time doesn't exist at that time in its entirety but exists there by having a present temporal part. Its other temporal parts, following perdurantism, exist at other times but, here comes the presentist's claim, those other times don't exist. But why claim that objects have temporal parts at other times than the present if these parts don't exist? According to Brogaard (2000), this is the best way for perdurantism to avoid what is, according to her, the main charge against it in its non-presentist form: that it entails a changeless world. Indeed, it is considered by many as a serious objection to the non-presentist version of perdurantism (four-dimensionalism) that it entails the denial of change in the world. The objection goes as follows.

My neighbour Cyrano, at some past time t_1, has a big nose. Suppose, then, that he undergoes plastic surgery and so has, at a later time t_2, a small nose. What this amounts to, according to the four-dimensionalist, is that one of Cyrano's temporal parts has a big nose, and another has a small one. Thus, four-dimensionalists often take change to be very much like spatial variation (which echoes the eternalist's general claim that time is space-like), since change is accounted for as the having of different properties by different parts. But when considering this account of change, some object to it[1] by claiming that what we want to give an account of is how a single object, a single individual like my neighbour Cyrano, can change, and the four-dimensionalist is telling us a story about different objects (different temporal parts) having different properties, and this is not the story we wanted to be told. What we have is not change of an individual, but replacement of one changeless object (one temporal part) by another changeless one. Instead of saying that Cyrano changed from t_1 to t_2 from having a big nose to having a small one, the four-dimensionalist says that the t_1-part of Cyrano changelessly has the property of having a big nose and the t_2-part of Cyrano changelessly has the property of having a small one. Since Cyrano himself is unable to lose or gain any such properties, this

[1]For statements of this objection see, for instance, Brogaard (2000), Haslanger (2003), Sider (2001a), Simons (2000b). The origins of the objection seem to be found in McTaggart (1927).

is why there is no room for genuine change in the four-dimensionalist's world. So, no concrete particular can ever genuinely change.

Brogaard claims that the presentist variant of perdurantism is capable of avoiding this objection (if this were the case, it would also show that presentism and eternalism are different): the t_1-part of Cyrano comes out of existence, by the passage of time, while the t_2-part comes into existence, and is then, in turn, replaced by another temporal part, and so on—thus the four-dimensionalist view, that "[a perduring object like Cyrano] has temporal parts with different properties, just as a multicoloured strip of paper has spatial parts with different properties, and neither case involves change in the sense in which this word is commonly understood" (Brogaard 2000, p. 342), is avoided because there is no such 'strip'. And since only one temporal part of Cyrano exists, namely the present one, the only properties instantiated are the properties instantiated by it now—there is no having tenselessly any property and there is genuine change in the world, for, as she puts it "[…] the coming into existence of a new stage [i.e. temporal part] with a different non-relational property is a real change – and this in a way that captures our most basic intuitions according to which a change has taken place if the object stage that presently exists has different properties from those that existed previously" (Brogaard 2000, p. 348).

I do not see the force of this answer to the objection (and so, I cannot appeal to it to establish a difference between eternalism and presentism). Suppose that the 'no-change objection' succeeds against four-dimensionalism. The core idea of the objection is that, instead of having a case where an object genuinely changes in its intrinsic properties, we have a case where one object (one temporal part) is replaced by another. Peter Simons, for instance, makes this claim when he says that the "four-dimensional alternative is not an explanation of change but an elimination of it, since nothing survives the change which has the contrary properties" (Simons 2000b, p. 65). Now, if this objection applies to four-dimensionalism, it obviously applies to presentist perdurantism as well—what we have, in the situation as described by Brogaard, is not one and the same object that would change any of its intrinsic properties, but a series of numerically distinct objects coming into and going out of existence, when one of them is continuously replaced by *another*. The objection, then, applies here with as much force as before.

But perhaps the presentist perdurantist could claim that her view, but not the four-dimensionalist's, can accommodate the claim that there is change in what exists—and this is perhaps why Brogaard thinks that it can answer the no-change objection. Indeed, the four-dimensionalist's ontology is a static one, since all times, past, present and future, equally exist, while the presentist component of presentist perdurantism allows for a world where what exists changes, since only the present time is real. But let us be careful about what such a claim is about: what we have here is a difference in what exists, since the reality's stock contains, for example, Cyrano with a big nose at some time, and does not contain such an entity at a later time. But how does such a claim answer the no-change objection? It doesn't. It is true that the reality's total stock is different from time to time—but *such* a claim turns out to be true even under four-dimensionalism, since reality's stock *at some*

time is different from reality's stock *at some other time*. The only difference is that, under presentist perdurantism, the reality's stock at a time is the reality's stock *simpliciter*, but this could hardly provide an answer to the no-change objection—granted, the defender of such a view could claim that "reality's stock *simpliciter* changes", but what else could such a claim mean, except, as we have seen, that reality's stock is different from one time to another?—which, again, is true even under four-dimensionalism. Furthermore, and most importantly, even if there were a difference between the two views with respect to a 'change' in what exists (reality's stock *simpliciter*), there certainly is no relevant difference in the account that the two views provide of what we wanted to account for in the first place: intrinsic change of an entity such as Cyrano. Exactly as under four-dimensionalism, nothing (that is, *no one thing*) undergoes intrinsic change under presentist perdurantism—what we have in both cases, to repeat the objector's charge, is not change of an individual, but replacement of one changeless object (one temporal part) by another changeless one. And the same goes, of course, for the world as a whole (reality's stock at a time)—the world is simply replaced by another, with the passage of time. It seems to me, then, that if the no-change objection succeeds against four-dimensionalism, it succeeds against presentist perdurantism as well. The diagnostic here is, then, that what causes trouble, according to the objector, if one wants a good account of change, is not the 'eternalist half' of four-dimensionalism, but rather its 'other half': perdurantism. And so, it is not presentism that can save the case of four-dimensionalism. To yield a satisfactory and intuitive account of change, the objector would probably argue, the cure is not presentism, but endurantism. Thus, until now, none of this can play the role of establishing a significant difference between eternalism and presentism. But it does help to understand better the inner workings of the views at hand, which will now help us to see how they *are* different, by considering an independent difficulty with the presentist perdurantist view, that four-dimensionalism does *not* encounter.

Remember: presentist perdurantism claims that at the present time t_1 an object such as Cyrano doesn't exist in its entirety but exists there by having a t_1-part. The perdurantist component of this view would push us to say that he also has the rest of his temporal parts existing at other times, but according to presentism, those other times don't exist. But how is it possible to claim that material objects have temporal parts at other times than the present if these parts don't exist? Of course, following presentism, one could *say* that they *existed* and exist no longer, but in what sense would they be *parts* of the object? The very plausible principle involved here was put forward by Trenton Merricks: "an object cannot have another object as a part if that other object does not exist" (Merricks 1995, p. 524).

According to Sally Haslanger, this is in no way problematic to the holder of presentist perdurantism (see Haslanger 2003, p. 11). Her grandmother, says she, is part of her family even though she does not presently exist, so if her family can have a non-existent part, why couldn't Cyrano? But such a line of 'argument by analogy' does not seem to be of great support, since typically a family and a material object like Cyrano or a table are conceived as different kinds of entities; thus, they are not analogous cases, and so any argument based on an alleged

analogy is misguided. In order to make this argument by analogy sound and persuasive, it would be necessary first to show that a family is best conceived of as a material object like a table, but Haslanger does not do that—and the burden of proof *is* on her, since one could very well plausibly argue that a deceased member of a family is not a *part* of it: a family is probably best conceived of as a plurality, like a football team, and exactly as a football team can lose one of its members when this member ceases to exist, a family can lose a member in the same way, and in both cases the lost member is not a part of the team or the family anymore. So Haslanger's example doesn't *prima facie* seem to be a good one, since the relation that family members or football team members bear to families and teams is a different relation from the parthood relation, and so this example cannot establish here that any non-existent object could be a part of anything existent, in the strong sense of "part" required by perdurantism—a doctrine according to which ordinary objects like tables are made up of temporal parts.

Concerns about family members aside, the main ontological difficulty here is that it seems very hard to admit that the objects (temporal parts) that compose another object (the whole Cyrano) exist only one after another, and so fail to ever make up the whole as they should. Cyrano is supposed to be an aggregate of his temporal parts, but there never is a time (or time-span) at which such an aggregate exists.

Lombard (1999) thinks otherwise. As he points out, rightly, one must carefully distinguish between two senses of "exist" if one is a perdurantist (both presentist and eternalist). First, there is a straightforward sense in which instantaneous temporal parts (let us admit here that there are such things, even if the perdurantist is not committed to them) exist at a certain time—if such entities exist at a certain time, they exist at this time entirely (they are three-dimensional entities) and they have all of their (spatial) parts at this time. Second, there is a derivative sense in which Cyrano, a whole composed of all of his temporal parts, exists at some time t—in this sense Cyrano exists at t in virtue of having a temporal part that does; but one is enough, he does not need to have all of his parts at t. Of course, it is the second, derivative, sense that is the interesting one for the perdurantist here, the first one being accepted by everyone: if there are any three-dimensional, instantaneous entities, it is uncontroversial that they exist entirely at the time they do.

Criticising Merricks's claim that an object cannot have another object as a part if that other object does not exist, Lombard says that "what is obvious is only that an object that exists *at a time t*, cannot have, *at t*, another object as a part, if that other part does not exist *at t*. But what the perdurantist wishes to say is *not* inconsistent with that. […] What exists now in [the derivative] sense – [Cyrano] – is something that does (at some time or other) have parts that do not exist now; but what exists now in that sense does not now have those parts" (Lombard 1999, p. 256).

But let us consider a true statement like "Cyrano has a present temporal part, but he is not identical to it". The problem here is simple: what is the referent of "Cyrano" and "he" in this statement? That is, what is this allegedly existent object

that we are making reference to by these words? Does this object exist? Of course, it doesn't—only a part of it does. Perhaps, the presentist perdurantist would say that the other parts *existed* and *will exist* and that there is a sense in which we can *speak* about Cyrano's being composed of all his parts—but such a strategy does not seem to be available here, since it would mean that one is taking seriously something like an extra-temporal standpoint from which one refers to an entity composed of different temporal parts existing at different times. The eternalist could do that, of course, but not the presentist, since doing this would be like considering the different times as equally real. It seems that the only thing the presentist can do is to see things from the standpoint of some determinate moment of time (the present) and, from this point of view, nothing that could be the referent of "Cyrano" is available. So, in what sense can the referent of "Cyrano" be said to exist? In reply, Lombard would probably say that it exists derivatively—but what does this mean here? In the presentist's vocabulary, the 'normal' meaning of "exists" is "exists now"—only what exists at the present time 'really' exists; remember that presentism is a doctrine about what there is in reality's stock, and that the doctrine claims that there is nothing more than the presently existing things. But now, the presentist perdurantist is telling us that there is more—that there is another, derivative, notion of existence according to which things composed of non-present (non-existent) things exist. But, first, it seems to be a strong departure from one of the central claims of presentism to introduce two concepts of existence—one that sticks to the presentist view, and another that does not seem to; and second, those two senses of "exist" are really distinct and irreducible one to the other. Compare to the case of the four-dimensionalist: she also uses two senses of existence, the 'ordinary' one, and the derivative one, but here, the derivative sense does not carry any new ontological commitments—it only tells us that something can exist at a certain time by having a temporal part here, but it does not involve anything more than what there already is in the first, non-derivative, sense of existence—so here, the derivative sense of existence is only a device to accommodate ordinary language, but nothing more.

So, it seems that the notion of having non-existent parts carries with it an ill-motivated plurality of notions of existence. But even if such notions of existence were to be accepted and endorsed, this would not leave the presentist perdurantist view free of problems concerning the having of parts that don't exist.

To see this, let us make a small detour and first consider another rescue mission that the defender of the presentist perdurantist view might want to undertake to answer the problem we had: that the temporal parts that compose an object exist only one after another, and so fail to ever make up the whole—and so it seems that Cyrano never really exists. Here is a remedy that is readily at hand: deny that Cyrano is a four-dimensional whole made up of temporal parts, and claim, rather, that he *is an instantaneous temporal part* which persists through time by having other temporal parts at other times as *temporal counterparts*. This amounts to a rejection of the more traditional perdurantist 'worm view' in favour of the so-called 'stage view'—since ordinary objects like Cyrano, according to this view, are the instantaneous stages rather than the worms made up of them. (This is the view

defended by Sider (2001a)). It is not my purpose to discuss the stage view here, I only wish to see how relevant it is to the combination of perdurantism and presentism. And it is obvious that it has the nice advantage to answer our objection: if Cyrano is an instantaneous stage rather than a temporally extended worm, then there is, of course, no problem about having non-existent parts, since nobody claims that he has any, and there is no problem about how successive stages could make up a whole, since nobody really cares about the wholes (Sider claims that the wholes exist in addition to stages, but that these are not the ordinary objects we usually care about and quantify over—so at least the pressure on the presentist perdurantist becomes much weaker here).

So isn't there a good reason for the presentist perdurantist to become a stage theorist? I think not. Consider the claim that Cyrano now has a small nose, but he had a big nose before. The stage view provides a counterpart-theoretic analysis of such a claim: Cyrano now has a small nose, but he has a past counterpart that has (had) a big nose. Now, what is needed for Cyrano to have such a counterpart? Two stages are counterparts iff they are related by the counterpart relation. The counterpart relation is a relation of similarity, some sort of spatio-temporal contiguity and/or continuity, and causality. Actually, no stage theorist (including Sider) says what exactly the nature of the counterpart relation is, but my point here is simply that whatever the counterpart relation is, it just cannot hold between different stages if one is a stage view theorist who wants to be a presentist as well. Take Cyrano at t_1 with a big nose and Cyrano at t_2 (the present time) with a small nose. These two different individuals are supposed to be counterpart-related. But how could they ever be? How could a non-existent individual (Cyrano at t_1) bear any degree of resemblance and have any other (spatio-temporal and causal) relations to an existent flesh-and-blood individual (Cyrano at t_2—the present time)? Nothing non-existent is sufficiently similar and related to anything existent to be counterpart-related (if it makes sense at all to even speak about 'non-existent things'). And generally, the counterpart relation will never hold between the two individuals (the two numerically distinct Cyranos) simply because there never is a time when the two individuals both exist—and so there never is a time when both relata of the counterpart relation exist. How, then, could the counterpart relation ever succeed in doing the job it promises if the relata that it is supposed to relate never both exist? Of course, one could say here that two individuals are counterparts iff, if they were both present (that is, if they both existed), then they would be counterpart-related, but such a situation never is the case, and so the conditional here would always be vacuous.[2] In short, then, the stage view does not really help the business of the presentist perdurantist because, even if it seems to answer the objection about parts that don't exist, it immediately yields a different but parallel objection about counterparts that don't exist.

And it is easy to see how this problem also makes trouble for the presentist perdurantist who wishes to maintain the worm view—exactly as different

[2]This parallels what Lewis (1986, p. 238) says about modal counterparts.

counterparts need to be related by a counterpart relation in order to be counterparts, different parts of four-dimensional worms need to be 'glued together' in some way in order to make up the wholes that are the individuals we are interested in, like Cyrano. Finding such a glue (that is, a unification relation that makes the successive temporal parts of a single four-dimensional worm ontologically stick together) is not an easy task even for the four-dimensionalist, but for the presentist perdurantist, the task just seems impossible to be carried out. For what would such a gluing relation be? Again, it might involve resemblance, or causality, or spatio-temporal contiguity, or something else—in fact, whatever features serve the stage theorist to load his counterpart relation can also serve the worm theorist as being the glue. And so, of course, the same problems as those we have just seen with the stage view will appear for the worm view: how could one existent thing and one non-existent thing be glued together (if, again, I may be allowed to even say such a weird sentence)? That is, what kind of ontological glue would be needed in order to make it the case that mereological composition takes place between a thing that exists and nothing? Perhaps one could propose, as a remedy, to follow the line of almost all four-dimensionalist's who are friends of the principle of an entirely unrestricted mereological composition (for independent reasons, mainly to avoid problems with ontological vagueness)—so that the glue relation might not be restricted at all. But however unrestricted, it certainly cannot be *that* unrestricted—*un*restricted mereological composition is *restricted* to existent things only, and any attempts to take away even this restriction would lead one to weird places where no sensible metaphysician (I hope) wants to go—like commitments to individuals made up of the top half of Cyrano's body today, and all of the tropical fish of the 19th century, *and* two fire-breathing dragons.

§9. In this chapter, I have considered the claim that eternalism and presentism are equivalent. I have discussed the strategy of alternative languages and interpretative charity, and found it too general, providing equivalence claims too cheaply and too easily. One needs to pay careful attention to the inner workings of the two views in order to see how they work and how they do the job they are designed for, before being able to make any equivalence claims. We have seen that presentism, *contra* eternalism, is not compatible with perdurantism, or at least that such a combination leads to serious difficulties (not raised by the combination of eternalism and perdurantism). If this is so, then *by doing first-level metaphysics*, we have established a meta-metaphysical result: presentism and eternalism are different, since they have different implications with respect to some connected metaphysical issues, that is, they do not play the same theoretical role in the same way, at least in some cases.

Chapter 4
The Importance of Being Primitive

Abstract This chapter concerns primitiveness. Indeed, my way of arguing for or against equivalence claims in the previous chapters largely depends on the nature of primitives and on the role they play in each of the theories involved, and it is now time to be explicit about this. In general, in most metaphysical debates a lot depends on primitives—indeed, metaphysical theories rely heavily on the use of the primitives that they typically appeal to. So, I will emphasize here the utmost importance of primitives in the construction of metaphysical theories and in the subsequent evaluation of them. I will claim that almost all of the explanatory power of metaphysical theories comes from their primitives, and I will scrutinize the notion of "power" and "explanatory". All together, these points will naturally lead me to defend a global view on the nature of the metaphysical enterprise: what is at stake in metaphysics is to find out not just what there is or what there is not, but what is more fundamental than what—to find out what are the best primitives. Relationships between my view and the current debate concerning the notion of grounding will be discussed.

§1. In Chaps. 1, 2, and 3, I examined some cases of metaphysical equivalence and non-equivalence, and I defended a methodological and metaontological view which emphasizes that when asking the question "Are metaphysical debates substantive or verbal?" the correct answer is "It depends." Some debates are substantive, some debates are merely verbal, while there is room for the view that a debate is partly merely verbal but not entirely and that some parts of it are substantive and decidable by philosophical methods. As we have seen, it is possible, and it is the case with respect to the persistence debate (Chap. 2), that, inside a debate, some points are merely verbal while other are places of substantive disagreement. The methodological point I insisted upon is this: the best way to do meta-metaphysics is to do first-level metaphysics.

This Chapter concerns primitiveness. Indeed, my way of arguing for or against equivalence claims largely depends on the nature of primitives and on the role they play in each of the theories involved, and it is now time to be explicit about this. In general, in most metaphysical debates a lot depends on primitives (as we will also see in Chap. 5)—indeed, metaphysical theories rely heavily on the use of the primitives that they typically appeal to. So, I will emphasize here the utmost importance of

J. Benovsky, *Meta-metaphysics*, Synthese Library 374,
DOI 10.1007/978-3-319-25334-3_4

primitives in the construction of metaphysical theories and in the subsequent evaluation of them. I will claim that almost all of the explanatory power of metaphysical theories comes from their primitives, and I will scrutinize the notion of "power" and "explanatory". All together, these points will naturally lead me to defend a global view on the nature of the metaphysical enterprise: what is at stake in metaphysics is to find out not just what there is or what there is not, but what is more fundamental than what —to find out what are the best primitives. Relationships between my view and the current debate concerning the notion of grounding will be discussed.

§2. In Chap. 1, I argued in favour of some equivalence claims, as for example the equivalence between BTT (Bundle Theory with Tropes) and STT (Substratum Theory with Tropes). My argument was roughly the following: (i) BTT and STT are equivalent; because (ii) the unifying device called "substratum" in STT and the unifying device called "compresence" in BTT are identical (metaphysical equivalence); because (iii) they play the same theoretical role in the same way; and (iv) they are theoretical entities (that is, they are individuated by their theoretical role).

But perhaps some will not be comfortable with (iv) because they will feel that the substratum and the compresence relation, even though they do the same theoretical work, are not 'just' theoretical entities but really *are* metaphysically *different* things. Or perhaps some will not be happy with (iii) and (iv) being enough to justify (ii), and consequently to justify (i), because they believe that playing the same theoretical role is not enough to justify the claim that there is a *metaphysical* equivalence. Depending on how one takes these worries, one might be tempted to accept one or the other of the following conclusions:

> *Strong Conclusion*: BTT and STT are *metaphysically equivalent*.
>
> Argument: compresence and substratum are theoretical entities, which means that they are individuated by their theoretical role. Since the theoretical role they play is the same, they are the same theoretical entity.
>
> *Weak Conclusion*: it is *epistemically under-determined* which one of BTT or STT we should choose.
>
> Argument: compresence and substratum are metaphysically different entities, but they play the same theoretical role in the same way, and STT and BTT have the same explanatory power (as far as we metaphysicians are concerned, they both do the job we want them to do).

This example concerns BTT and STT, but we have also seen other similar cases, especially relationism and substantivalism about time, or BTU and STU (Chap. 1). We have seen how both sides of the debate use *problem-solvers* in the same places, and so primitively solve a problem in the same—primitive—way. The Strong Conclusion, then, insists on the idea that these problem-solving primitives are literally numerically identical entities: they are one and the very same thing, since they are theoretical entities that are introduced by the metaphysician to do a job, and consequently they are *individuated by their theoretical role*. If, then, this role is the same, they just turn out to be one and the same thing under different names. As we have seen, the (Strong) conclusion to be drawn from this is that it becomes really hard to see what difference there is between relationism and substantivalism, between BTT and STT, and between BTU and STU, since both sides of each debate use one and the

same primitive problem-solver in many crucial places and they are also otherwise structurally extremely similar—thus, it is true for each of these three pairs of views that the two allegedly competing sides are actually metaphysically equivalent.

But one may think that playing the same theoretical role is not enough to justify *metaphysical* equivalence. On this view, it is correct to draw a Weak Conclusion of *theoretical* equivalence, that is, equivalence with respect to what the two allegedly competing theories can *do* and with respect to how they face their theoretical challenges, but this does not extend to the stronger claim about what they *are*. A substratum or a substance-instant, this weaker view insists, just is not the same thing as a bundling relation or a relation of simultaneity, even if it does the same job within the theory. It is individuated by its nature, not just by its theoretical role, and its nature is not the same. Thus the Weak Conclusion claims that the two sides of the debate are not metaphysically equivalent but that it is *epistemically under-determined* which one we should choose since they both do the same job in the same way.

I think that the Strong reading of my equivalence claims is superior to the Weak one. The Strong Conclusion takes very seriously the *functional role* that problem-solvers play in the theory. By its very nature, a primitive being primitive, it is non-analysable and we are not really given any information concerning its nature; we are told *what it does* rather than *what it is*. So it is what it does that counts—after all, that's what any primitive is introduced for in a theory in the first place (otherwise there would be little justification for having it). Thus, primitives are individuated by what they do, what their functional role in a theory is, and, as a consequence, two primitives that do the same job just turn out to be equivalent for all theoretical purposes and metaphysically equivalent as well: they just are one and the same thing referred to in two different ways. We have seen in Chap. 1 that with respect to some traditional alleged points of disagreement, the views I have examined contain a primitive machinery that does the same job at the same place in the same way (that is, in a primitive way).

This is then what one can call a "functional view" of primitives, and this is the view that I have been implicitly using throughout the discussion in Chap. 1, and that I believe to be correct. The alternative (Weak) view, that I will call the "content view" of primitives, claims that not only do primitives have a function in a theory but also that they have a nature (a content). For instance, in the case of the Bundle Theory and the Substratum Theory this view insists that the unifying device involved in the former is a *relation* while the unifying device involved in the latter is a *substratum* and this just is not the same thing, however similarly they may work.

If the functional view is correct, the primitives used by the two sides of a debate just turn out to be the same thing and given the similarity of structure and equivalent explanatory power of the two allegedly competing theories, it appears that the difference between the two sides is no more than terminological. I leave it, then, an open question whether *this* terminological difference can be a good reason to prefer one side over the other Hirsch (2005, 2007, 2008), when making an equivalence claim about the debate between perdurantists and endurantists, insists that, while there is only terminological difference between the two sides of the debate, the endurantist language is closer to ordinary language and so should be preferred.

Alternatively, one can see closeness of theoretical terminology to ordinary language as irrelevant and simply claim that it does not matter at all which side of the debate one chooses to embrace—or better, that we should simply refrain from choosing. This question about which is the better terminology left aside, I would like to insist on the following: in all theories we have seen in Chap. 1, and in metaphysical theories in general, the theories' primitives do a *big* part of the job—indeed, without its primitives none of the views we have examined would even begin to work and primitives are used in every crucial place where a serious problem needs to be solved or an important phenomenon (like persistence, unity,…) is accounted for (in Chap. 5, I will discuss in detail more examples of this). Under the functional view, it is then no wonder that if the primitives turn out to be equivalent (since they are individuated by their theoretical role and the role is equivalent) then the theories that contain them appear to be no more than terminological variants.

§3. Let me say a bit more about why I think that the functional (Strong) view is to be preferred. Take the case of the Bundle Theory and Substratum Theory: the friend of the content view will insist that one side's problem-solver is a *relation* and the other side's problem-solver is a *substratum* and so they are entities with a *different nature*. Now, I think there are two possibilities: either any difference that this can make will be a functional difference, or this is just stubbornly sticking to terminology. For instance, the content view might say that a substratum is 'ontologically independent', that is, it can exist without exemplifying any properties, while the relation of compresence cannot just 'be there' and relate nothing. But if that were to be a way to claim that there is a difference between the two primitives, then it would be a functional difference: there is something that one primitive can *do* (standing alone, not unifying anything) and that the other primitive cannot *do*. So this is not going to give the friend of the content view what she needs. But it could, of course, block any equivalence claim since this would actually show that the two unifying devices do *not* always play the same theoretical role in the same way, and are thus not equivalent—or at least the equivalence claim has to be restricted to some cases only, but cannot be generalized. This is, of course, something that I am open to: if it can be shown that the two unifying devices do not have exactly the same function, any of the (Weak or Strong) Conclusions could only be partial. In Chap. 1 §14, I discussed the case I just mentioned about 'independence' with the result that there is actually no difference with respect to the two primitives, but in principle it is an open possibility that a place where they do play a different role can be found (but until then, the equivalence claim holds). Anyway, even if this were the case, the difference between the two primitives would be a functional difference so such a case would count in favour of the functional view. In principle, we should always expect any difference between primitives to be functional, with no surprise: since primitives are introduced in the theory by the metaphysician who needs them because she cannot make the theory work without them, she'll typically always introduce primitives to do a theoretical job, otherwise she would not even bother with them in the first place —this is why it seems to me quite obvious that any difference we could find between

primitives will turn out to be functional.[1] If someone insisted that on the top of the functions they play in the theory, primitives have a non-functional content, this would then amount to insisting that she postulates *a difference that makes no difference*—and I can only say that I see little reason for doing anything like that when building a metaphysical theory. Such an attitude towards primitives would be having an unreasonable soft spot for the words one uses—words like "substance", "substratum", "relation", and so on.

Perhaps the friend of the content view could argue that even if two primitives play the same theoretical role—that is, they have the same function—they still have a different nature and a non-functional way to see that is to embrace one of the two primitives as being more *intuitive* than the other. Something like this might have played an important role in the debate between substratum theorists and bundle theorists where the latter sometimes, in this vein, characterized the substratum as a mysterious entity. So can *intuitions* help us to distinguish between two primitives that have the same function? I do not think so: there are some worries that arise here, showing that intuitions are not really pertinent in this matter. Firstly, intuitions do not seem to be relevant in the field of basic metaphysics which is just too abstract and theoretical for any useful intuitions to arise—unlike in other less basic debates; for instance, imaginary cases or Star Trek stories of duplication of persons in the debate about personal identity allow us to give rise to some carefully formulated and useful intuitions that can probably do some helpful work in the understating of our concept of a person and its conditions of persistence through time. But when fundamental and highly abstract metaphysical issues are concerned, such as those that we are concerned with in our case of the theories discussed in Chap. 1, there just do not seem to be any useful intuitions around: these are not matters where anybody can have any intuitions, except, again, misleading intuitions that come from attachment to *words* like "substance" or "relation". Secondly, all intuitions suffer from being too unsettled and variable from one thinker to another and even over time for one and the same thinker, and there are conflicting intuitions, good and bad intuitions, as well as weak and strong intuitions—thus, they do not seem to be a very reliable guide (more on intuitions in Chaps. 5 and 6).

I think that this leaves us with the functional view, and with the consequence that if my arguments above are correct, a Strong Conclusion can be drawn. To sum up some central points of my approach: (i) theories such as the ones discussed in

[1]This is also the reason why I prefer to restrict my claim that the functional view is correct only to the case of *primitives*, and not all other cases of functionally equivalent non-primitive entities or bits of theories. For instance, you might think of the case of two properties that have each a certain function in a world W and that 'switch' their function in a world W' while each keeping its identity: in such a case, it seems that something like the content view is to be preferred, with respect to the nature of properties. The difference between such cases and the case of primitives is that primitives are theoretical postulates that are introduced in a theory by a metaphysician who needs them to perform a certain job, a function, and so this is why I think there that the functional view holds, while properties are not just theoretical postulates and so it *may* very well be that with respect to properties the content view is more adequate since it seems at least *prima facie* more plausible that they have a nature that is not reducible only to their function (But of course, the opposite may turn out to be true; all I want to say is that I wish here to restrict my claims only to the case of primitives.).

Chap. 1 (and those we will see in Chap. 5) only work thanks to the job done by their primitives; this is, of course, to be expected and perfectly in order since otherwise there would be no reason to introduce the primitives in the first place; (ii) thus, it is crucial to recognize that, indeed, these primitives are introduced by the theorist, they are theoretical postulates that are introduced to do a job: they are there for what they can do, they are functions within the theory; (iii) as such, they are individuated by their functional role; this has the consequence that if two primitives have the same function they turn out to be not only theoretically but also metaphysically equivalent.

§4. One might worry here whether the method I recommend in order to see whether two theories are equivalent or not—that is, to look at how the theories and their primitives *work*, how they *function*, to see whether any equivalence claim can be drawn—does not give us equivalence too easily and too cheaply. Indeed, in a very general way one could use my claims to say that whenever two theories or two primitives explain the same thing, they turn out to be equivalent. The materialist and the dualist both account, in their own terms, for the fact that I feel back pain this morning and for the having of qualia in general—does it then mean that I would say that they are equivalent, since they accomplish the same work in the end? Do their primitives "have the same function" because they somehow play the overall very general role of "helping the theory to solve the mind-body problem"? Doesn't, then, my functional view of primitives make claims of theoretical equivalence between them, and, consequently, equivalence between theories that use them, too cheap?

In order to claim that two theories or two primitives are equivalent, not only must they do the same *overall* job, but they must also do it *in the same way* in the sense that is relevant to an apt *level of analysis*. Whether an equivalence claim is too cheap or not depends, then, on the level of detail the claim provides. If such a claim is too general, it is very cheap; but if it is elaborated in detail and shows how the inner workings of two theories are similar and how their primitives behave in the same way at the same crucial places in the theories, then it is worth your money. Problem-solvers are here to solve problems, and so it is appropriate to evaluate what they are and how they function relative to a specified problem (or, set of problems)—after all, this problem is the very reason for postulating them in the first place. Now, if the problem is specified by saying "we need to solve the mind-body problem" then the level at which the problem-solvers that are involved in its solving should be evaluated and compared is somehow general. Thus, the answer to the question whether two primitives "do the same job in the same way" (= have the same function) is relative to a way of specifying the problem to be solved by these problem-solvers. Similarly, the answer of whether two *theories* are equivalent is also relative to a set of problems, puzzle cases,… that we want them to provide a treatment for (after all, that's what theories are for). It will not do to specify these problems in a *too general* way or in a *too detailed* (i.e. question-begging) way. I think that there usually is a "best" or "preferred" or "non-question-begging" or "appropriate" way to put the problem to be solved, and it is part of the metaphysician's job to identify what is the most apt/relevant/best/appropriate level of analysis, and then see how primitives (and theories) behave relative to it. Some levels are too general to be of any real use, some are too specific to the point that they become question-begging, and our job as

metaphysicians is to determine what the correct/useful/best/appropriate/… level is. There is no general, principled recipe for doing this; it is not a meta-metaphysical or purely methodological matter, rather it should be carefully done in detail when one does first-level metaphysics and when one tries to make the best sense of a metaphysical problem and the theories that try to solve it.

§5. Bearing the functional view of primitives in mind, and focusing on the fact that they do most, if not all, of the theoretical jobs done by our theories (again, I will say more about this in Chap. 5), one may start to ask itchy questions about the notion of 'explanatory power'. Where does the explanatory power of our theories come from? If what precedes is right, is all or almost all of their explanatory power just primitively postulated? What and how exactly do such theories explain? How can a primitive explain anything? After itchy questions, one may want to start to ask sceptical ones: What are such theories good for? Aren't they just clever and elaborated ways of answering our questions by primitively postulated 'answers'? If so, what have we gained by building such theories?

I am not a sceptic, but I do feel itchy. In what follows I will discuss the notion of explanation and explanatory power, and try to answer some of these uncomfortable questions.

§6. Among the different types of explanations, *why-explanations* hold an important place: "Why did dinosaurs die out? Because a giant meteor collided with the Earth." These explanations are often causal explanations. But this is not what we're looking for here. Closer to home, here is a kind of explanation that it will be useful to consider in some detail; this is an example familiar from the literature about the mind-body problem: "Why does lightning occur just when there is an electric discharge between clouds or between clouds and the ground? Because lightning simply *is* an electric discharge involving clouds and the ground. There is here only one phenomenon, not two that are correlated with each other; and what we thought were distinct correlated phenomena run out to be one and the same. Here the apparent correlation is understood as *identity*" (Kim 2006, p. 85).

The relation between the explanadum and the explanans is simply identity here. Take the example of the substratum theory and Max Black's universe with the two perfect spheres: the explanandum is the numerical diversity of the spheres, the explanans is the having of a numerically different substratum. It does then seem like the right thing to say, if you are a friend of substrata, that the numerical diversity of spheres S1 and S2 *just is* or *consists in* their having a numerically different substratum. *Prima facie*, it seems then that the relation between a primitive problem-solver and the phenomenon it explains is identity. But there are some problems with this view.

First, it is not true that all identities are explanatory, as, for instance, Ruben (1990, p. 219) argues. To take the example of lightning, the identity "lightning = lightning" is not explanatory, while the identity "lightning = atmospheric electric discharge" is, because even if there is only one phenomenon involved in the case of the latter identity, it is conceptualized in two different ways. This teaches us that explanation is—unlike identity—an irreflexive relation.

Furthermore, one can offer reasons to think that even the identity "lightning = atmospheric electric discharge" is *not* explanatory. If the relation between

'substratum' and 'numerical diversity' (or, similarly, for compresence under the bundle theory) is identity, how does this *explain* anything? This worry is also familiar from the discussion of the mind-body problem. Kim (2006, pp. 97–98), when discussing the psychoneural identity theory, says: "[…] Our conclusion, therefore, has to be that both forms of the explanatory argument are open to serious difficulties. Their fundamental weakness lies in a problematic understanding of the role of identities in explanation, an important topic that has not received much attention in the literature. The only clear (and also simple) view is that identities function simply as rewrite rules in explanatory derivations–or any derivation, for that matter. […] We do not have to say that identities have no role to play *in* explanations. For they can help *justify* explanatory claims–the claim that we have explained something. […] It is only that identities do not generate explanations on their own."

The trouble, as I understand it, is that since the primitive problem-solver (the explanans) is actually the very same thing as the phenomenon we sought to understand (the explanandum), it is not very clear what we have gained by such an explanation; that is, by providing an explanation of what we wanted to understand in terms of a primitive that's actually the same thing as what we yearned to have a better understanding of.

One can respond to these worries by arguing that the relation of explanation is a lot like identity but that it is not identity. It is irreflexive, as we have already seen, but it is also asymmetrical. Granted, lightning *is* atmospheric electric discharge, and the phenomenon of numerical diversity *is* (say) the phenomenon of having different substrata. However, the explanation here does not consist just in pointing out this fact, it also points out the fact that the explanans is *more fundamental* than the explanandum. This is, then, what we gain, this is what we learn from a good explanation. Explanations of this type are such that one of the two sides of the explanation relation is more fundamental than the other (and thus, explanation is not only irreflexive but also asymmetrical). What we have here then is a sort of *explanation-by-identity*—but not *identity*—that parallels a growingly familiar notion of grounding: as Bricker (2006), Derosset (2010), Schaffer (2009), and Schaffer (manuscript) claim, if a is grounded in b, a is nothing over and above b. a, in other words, is an "ontological free lunch" in Armstrong's (1997) sense; the "ontological price", to use Schaffer's (manuscript) term, you pay for a and b is just whatever you would pay for b alone. Only in this sense can one talk about identity between a and b. (This kind of explanation goes around under many different names like 'in virtue of', 'just is', 'is grounded in', 'is', or 'consists in'—but these terms are tricky and are not always intended to mean anything like explanation. For instance, 'consists in' is sometimes taken to mean 'constitutes', which is a 'grounding' relation familiar from the debate about material constitution (Wiggins 1968; Heller 1990, 2008; Burke 1994; Bennett 2004)).

We encounter this kind of explanation-by-identity in very many cases of the workings of our metaphysical theories. For example, two particulars a and b's having the same property *is* their instantiating the same universal (if you are a friend of universals) where the latter is taken to be more fundamental than the former, or

a's persisting through time *is a*'s having temporal parts at different times (if you are a friend of temporal parts) where again the latter is taken to be a more fundamental phenomenon than the former. More often, one uses the locution "in virtue of": *a* persists through time *in virtue of* having temporal parts at different times.

In the case of a lightning and an atmospheric electric discharge, the chain of explanation goes on until the most fundamental level is reached—and what's the most fundamental level depends here on the current state of physics, it is largely an empirical matter. Not so, of course, in the case of explanation relations between primitive problem-solvers and explananda of our metaphysical theories. As metaphysicians, we are most typically not 'stopped' by empirical matters, but rather we find the end of our metaphysical discovery when we step on a concept that is not further analyzable without circularity—a substratum, a compresence relation, non-relational instantiation, resemblance, and so on, depending on your favourite metaphysical theory. Such notions are then taken by our theories to be too fundamental to be usefully (in a non-circular way) further explained. This raises an interesting problem.

Explanatory power is one of the main criteria we use to evaluate our metaphysical theories. After all, the very point of building a metaphysical theory in the first place is to provide an explanation for some phenomena that we want to better understand (particularity of particulars, sharing the same property, persistence through time, …). If what I said above is correct, and very close to what Schaffer (2009) argues for in a different way, the picture one gets of what metaphysics does and what it should do is thus not just to tell us *what there is* but, more importantly, to tell us, as Schaffer puts it "what grounds what", that is, in my way of putting it, to discover what are the most fundamental notions, which are primitive and which are not. The idea here is that metaphysics does not and should not give us a *list* that is a sort of inventory of what there is, but rather a top-bottom structure of relations of 'grounding' or 'explanation' between types of entities or between concepts, saying which are primitive, which are not, and which are more fundamental than others.

But now, the interesting problem arises from the fact that some such structures (some metaphysical theories) take as a primitive problem-solving explanans what its opponent takes as being the explanandum, and vice versa. For example, the theory of universals says that *a* resembles *b* because *a* and *b* both instantiate the numerically same universal. *a*'s resembling *b* consists in *a* and *b*'s instantiating the same universal. *a* and *b* resemble each other in virtue of instantiating the same universal. And so on. The phenomenon of resemblance *is* the phenomenon of instantiating the same universal, where the latter is more fundamental than the former. Identity (instantiating the numerically *same* universal) is fundamental, resemblance is not—it is *derived* from identity. But not so for the friend of Resemblance Nominalism: under her view, *a* and *b* have the same property because they are both members of the same resemblance class. Here, *resemblance* is the fundamental notion, and identity (sameness of properties) is derived.

How to tell then which one of these structures is better? How to choose between such competing metaphysical theories? In some cases, like those we have seen in Chap. 1, we do not have to make a choice: these are the cases where we discover

that the allegedly competing sides of the debate are actually not very different from each other—that indeed they are theoretically and/or metaphysically equivalent. But in many cases, this is not so, like in the case of the debate about persistence (endurantism vs. perdurantism—see Chap. 2), or the case of the theory of universals and nominalism (Chap. 5 will be largely devoted to this issue). Another way to ask the question is: how do we know that we are facing a primitive when we step on one? Is resemblance primitive? Is instantiation of a universal primitive? How do we tell which one of these lies at the bottom of the structure and which one does not? As we have just seen in the preceding paragraph, one theory can take as a primitive what the other takes as being higher in the hierarchy of its structure, and vice versa, and it works. Given that both sides of a debate have the same right to introduce their own primitive problem-solving devices when they need one, it seems very hard to find independent, objective grounds to choose one structure over the other.

I do not have a good answer to this question. In Part II, I will defend the view that the reasons to select one theory over its competitors are mainly grounded in aesthetic considerations, and in the evaluation of the aesthetic properties metaphysical theories possess. But let us wait with that until we get to Part II. What I would like to say now is that even if we do not have a nice and straightforward recipe to select the best structures, and to decide what is more fundamental than what, and which are the best primitives, we learn here to acknowledge the importance of these primitive problem-solvers, the importance of carefully thinking about the possibility of equivalence between them, and thus to recognize the utmost importance of the role that primitives play in (the building of, and evaluations of) our theories.

Thus we can answer these itchy questions: "What is at stake in the competition between metaphysical theories? What is this competition good for? What do we learn?" What is at stake is finding out which is the best primitive, and what is more fundamental than what—what explains what. Sometimes the answer will be a bit frustrating (while nevertheless being illuminating): these are the cases of theoretical and/or metaphysical equivalences between some allegedly competing views. Sometimes it may be argued that the answer is framework-relative, in a Carnapian way. And sometimes we will continue to fight tooth and nail to show, say, that resemblance is a better primitive than instantiation. Such work is not an easy one, and it cannot be done in a paper or in a book, rather it is the collective efforts of many metaphysicians that can bring any durable and good results. Such work, I suggest, is best seen as work on the primitives that sustain the structure of our metaphysical theories.

§7. In the first Part of this book I was interested in claims of equivalence, partial equivalence, and non-equivalence, as well as in methodological issues, and in the nature and the role of primitives involved in metaphysical theories. I found that the best way to do meta-metaphysics is to do first-level metaphysics, and that primitives should be the core of our interests (both when doing metaphysics and meta-metaphysics). As we will see in the next Chapter, primitives play a crucial role in most metaphysical theories, even in cases of clearly non-equivalent views.

In the next Part, I will meet the question briefly raised above: how do we choose between competing metaphysical theories? If it turns our that the two sides of a debate are equivalent, then we do not have to choose—that's one way of "choosing" to which Part I was devoted. But in the case of non-equivalent theories, such as presentism and eternalism (Chap. 3), or the theory of universals and resemblance nominalism (which I will use as my main example in Chap. 5), how do we say which is better? Read on, answers are coming!

Part II
How to Make a Choice Between Metaphysical Theories?

Chapter 5
How *not* to Make a Choice Between Metaphysical Theories

Abstract In Part I, we have seen various claims of metaphysical equivalence. The claim I make in this chapter, using the example of Trope Theory, the Theory of Universals, and Nominalism, is not one of them: I will not claim that these theories are in any sense equivalent. My claim will be different, yet related: the three views, while being different, do the same job in very much the same way when it comes to solving the problem of attribute agreement, and even when one acknowledges the differences there are between the competitors, there is little ground to pick a winner. I am not interested in this claim only for the sake of the debate between Trope Theory, Nominalism and the Theory of Universals. Indeed, as we shall see, the reasons why I claim that it is so difficult to choose one theory over the other can be applied to other cases as well. My claim here is mostly critical: I limit myself to showing that we metaphysicians are in an uncomfortable position with respect to these views when we try to choose between them, but I do not suggest a way out (this will be the task of the following chapter).

§1. Metaphysicians build theories. In order to do this, we usually first look at what others have built before. So, what happens is that, after having spent some time on a metaphysical problem or question, we end up with a list of theories that are candidates for solving or answering it with all their pros and cons—that is, all the arguments and objections that the collective efforts of all of us gave rise to. It's like having a *map* in front of us of all the ways to face the problem or the question. Put in this way, metaphysicians are involved in cartography of the logical space of metaphysical theories.

This is, of course, only a part of the metaphysician's work, for her purpose is not only to provide such a map but also to find the best route, that is, to select one theory as being better than the others. For my part, I find this to be an extremely difficult task, and that's what the second part of this book is all about: how to choose between competing metaphysical theories? The competing candidates are often of very equal forces and the choice between them is not an easy one to make. To be clearer, let us consider an example of a metaphysical debate, one which I think is among the most typical (and old) ones: the question of the nature of particular objects and their having their properties. In this chapter, I will look at the 'map' concerning this question, and consider various meta-theoretical criteria for evaluating and comparing the theories

J. Benovsky, *Meta-metaphysics*, Synthese Library 374,
DOI 10.1007/978-3-319-25334-3_5

on it. As we shall see, among the results we will obtain by doing this evaluative meta-theoretical work lies the interesting claim that once we examine carefully how these different rivals address the traditional problem of 'attribute agreement' (a problem that has been at the root of the very reason for developing these theories in the first place), we will realize that there is not much of a difference between the ways these theories handle the problem, and, in a more general way, we will then see that there is little reason for preferring one of these theories over the others.

In Part I, we have seen various claims of metaphysical equivalence. The claim I will make in this chapter about Trope Theory, the Theory of Universals, and Nominalism is *not* one of them: I will not claim that these theories are 'equivalent'. My claim will be weaker: the three views (which I shall articulate in more detail below), while being different, do the same job in very much the same way when it comes to solving the problem of attribute agreement, and even when one acknowledges the differences there are between the competitors, there is little ground to pick a winner.

I am not interested in this claim only for the sake of the debate between Trope Theory, Nominalism and the Theory of Universals. Indeed, as we shall see, the reasons why I will claim that it is so difficult to choose one theory over the other can be applied to other cases as well (where, possibly, they sometimes *might* lead to claims of metaphysical equivalence). In this chapter, my aim and my claim will be mostly critical: I will limit myself to showing that we metaphysicians are in an uncomfortable position with respect to these views when we try to choose between them, but I will *not* suggest a way out (this will be the task of Chap. 3).

§2. Let me say more precisely which are the theories I will be using as the case study in this chapter. The questions are: (1) What are material objects (such as tables, particles, lizards,…), what is their nature? (2) What are properties, what is their nature? (3) How do objects have their properties? I suggest here to consider views that answer all three questions, since if we took only one question at a time, we would never have theories 'complete enough' to be proper subjects of comparison and evaluation. A complete cartography of the various theories that provide answers to these questions would be a matter for several long books, but it is enough for my present concerns to use as an example a partial and quite schematic map, featuring what I take to be the three strongest, most representative, and typical views; namely

- Theory 1: Substrata that instantiate immanent universals
- Theory 2: Bundles of tropes
- Theory 3: Resemblance nominalism

Here is the schematic map:

Theory 1: Substrata that instantiate immanent universals

Answer to question 1	• Objects are bare particulars (substrata) that instantiate immanent (spatio-temporal, multiply locatable) universals
Answer to questions 2 and 3	• a is F iff a instantiates the universal F-ness • a and b are both F ('share the same property') iff a and b both instantiate the numerically same universal F-ness

(continued)

(continued)

Primitives, and other ontological commitments	• Bare particulars (substrata) • Immanent universals • Non-relational instantiation
Main objections (and replies)	• OBJ1: Infinite (Bradley-like) Regress (REPLY: the primitive instantiation is non-relational (Armstrong 1978)) • OBJ 2: Immanent universals are unacceptable because multiple location is absurd: "redness is at one meter from itself" (REPLY: this intuition was made for objects, not for properties (Lewis 1983; Loux 1998; Hawthorne and Cover 1998)) • OBJ 3: Identity conditions of substrata. In virtue of what is a substratum distinct from another? No attributes or properties can distinguish between them (Loux 1998) (REPLY: They *do* have properties. For what it is to have properties, according to the substratum theory, is to instantiate universals (Sider 2006)) • OBJ 4: Substrata are unknowable and they cannot be experienced. (Campbell 1981) (REPLY 1: let us concede that the introduction of bare substrata is incompatible with a rigorous empiricism, but […] the constraints the empiricist imposes on the metaphysical enterprise are unreasonably stringent. (Loux 1997, 1998), REPLY 2: Substrata *do* have properties, so the epistemological argument may be swiftly dispatched. We clearly can know what universals a thin particular instantiates, and so know what it is like; and in what other sense ought we be able to "know it"? (Sider 2006)) • OBJ 5: The idea of a bare entity with no essential attributes is incoherent, substrata have the property of not having any property (Loux 1997, 1998) (REPLY: only sparse properties count here, and 'having no property' is not sparse (Armstrong 1978; Sider 2006))

Theory 2: Bundles of tropes

Answer to question 1	• Objects are bundles of compresent non-repeatable properties (tropes)
Answer to questions 2 and 3	• a is F iff a has among its members/constituents/parts an F-trope • a and b are both F ('share the same property') iff a and b both have among their members/constituents/parts numerically different F-tropes that are exactly similar
Primitives, and other ontological commitments	• Tropes (their existence and nature) • Compresence • Exact resemblance of tropes; resemblance • Possibilia + counterpart theory
Main objections (and replies)	• OBJ1: variant of the problem with co-extensive properties (Manley 2002) (REPLY: use possibilia) • OBJ2: if F-ness is the set of all resembling tropes, since sets have their members necessarily, there could not be a single object that is F in addition to those that are there (Loux 1998) (REPLY: use possibilia)

(continued)

(continued)

	• OBJ3: the problem of naturalness (imperfect community difficulty) (REPLY: does not arise if we use *exact* resemblance (a primitive); more in Manley 2002) • OBJ 4: Infinite Regress—if we have exactly resembling tropes then the resemblance relations must resemble, and these must resemble, and so on. (REPLY: there is no regress thanks to the primitive fact that a and b resemble each other. We do not have two tropes and a relation of resemblance, but two tropes, and given these two, they resemble. (Simons 1994)) • OBJ 5: Problem with change. If an individual is identified with a bundle of properties, then if one of the properties changes, the bundle is not the same, and so, the individual who is the bundle is not the same. (Van Cleve 1985) (REPLY: use four-dimensionalism (Hawthorne and Cover 1998), or also endurantism (Benovsky 2008)) • OBJ 6: Modal analogue of the problem of change above. (REPLY: use modal counterpart theory) • OBJ 7: Properties are not the kind of entities that are capable of independent existence, they are not suitable to be ultimate constituents of reality (Armstrong 1997). (REPLY: this is no more than an undefended incredulous stare (Hawthorne and Cover 1998))

Theory 3: Resemblance nominalism, without paradigms

Answer to question 1	• Objects are not analyzed by this theory, they are taken as primitive
Answer to questions 2 and 3	• a is F iff a resembles all the Fs (the right-hand side being the more fundamental) • a and b are both F ('share the same property') iff they are both members of the same resemblance class
Primitives, and other ontological commitments	• Objects (answer to Question 1) • The fact that a and b resemble each other (but there is *not* a *relation* of resemblance between a and b; resemblance is a relational *fact* without there being such a relation) • Possibilia + counterpart theory
Main objections (and replies)	• OBJ1: the possibility that there is only one object that is F (Armstrong 1978) (REPLY: use possibilia (Rodriguez-Pereyra 2002)) • OBJ2: co-extensive properties, and the companionship difficulty Armstrong (1978), Campbell (1981), Jackson (1997), Manley (2002) (REPLY: use possibilia (Lewis 1986; Rodriguez-Pereyra 2002)) • OBJ3: the problem of naturalness (the imperfect community difficulty) Armstrong (1978), Manley (2002), (REPLY: Rodriguez-Pereyra (2002) using primitive resemblance) • OBJ5: Russell's regress (Russell 1912; Armstrong 1978) (REPLY: there is no regress Rodriguez-Pereyra (2002) using the primitive fact that a and b resemble each other)

(continued)

(continued)

	• OBJ6: resemblance nominalism and causality; causal powers of objects should depend on how objects are and not on how they are related to other objects (Armstrong 1978) (REPLY: causal powers *do* depend on how objects are, while objects are the way they are in virtue of resembling other objects—this is among the primitives of resemblance nominalism (Rodriguez-Pereyra, private communication)) • OBJ7: a resembles b because a is F and b is F, *not* the other way around (Armstrong 1978) (REPLY: this shows that on this point resemblance nominalism goes against intuition, but the view has theoretical virtues that outweigh this (Rodriguez-Pereyra 2002))

Of course, this schematic map is not complete and does not exhaust all the (variants of) arguments and objections there are for and against the three candidates. But all the main objections are there, and the map is complete enough to allow me to suggest the following:

- None of the objections listed is decisive, in the sense that none of the three candidates can be said to be *clearly eliminated* by it.
- Neither is one of the three candidates clearly eliminated by the *collective* force of the objections to it.
- More hazardously: it is likely, after centuries/decades of collective efforts of the community of metaphysicians, that no *radically new* objections will arise and clearly eliminate one candidate (but of course, it is possible).

What I want to suggest is that, unless a radically new treatment of these issues is found, it is *not* by making more objections or by raising the objections listed above that we will be able to choose one view over the others. Am I, then, saying that working on objections is futile? Of course not—working on arguments, objections and replies to them is a job of the utmost importance, since it allows two very important things: first, it eliminates candidates that do not work properly, and second, very importantly, it allows us to *articulate* the good candidates fully and in detail, by considering all the places where their claims lead as well as all the ways they can work and answer the question(s). This gives us a map that is complete in the sense that it contains all the good candidates, while not being unnecessarily overcrowded by having eliminated bad candidates, and is also detailed enough to provide sufficient information about the good ones. Working on objections, arguments, replies and proper ways of articulating theories is the highly important job of drawing the map, without which one would simply get lost. But once the map is drawn, something else is needed, as I think the example above shows: the objections and replies listed are simply not enough to clearly choose one view over the others.

This leads me to a second remark about the ‘map’ of my example, which connects it to the discussion from Chap. 4:

- The *primitives* of each candidate do a big part of the job, when it comes to answering the initial question(s) and replying to objections. It is only *if* we grant the views their primitives (and ontological commitments) that they can face all the objections.

I am not saying, of course, that the three candidates have all their objections easily answered and that they are problem-free, all I am saying is that all three views can plausibly face the objections that arise against them, with the important help of their primitives.

After all, everybody in the community of metaphysicians is aware of all these objections and replies, but there are still trope theorists, resemblance nominalists, or friends of universals, and they are rational and informed philosophers. If we all have the same data, and if we all work in a rational way, *why is it that we don't all end up defending the same view*? Put in this way, the question is perhaps a bit naïve, but I think that it is genuine, and I suggest that we keep it always in mind in what follows.

So, since first-level metaphysics (objections, arguments,…) does not suffice to make a choice among the candidates, and following the considerations above, I suggest now that we have a closer look at what meta-theoretical evaluative criteria there are for metaphysical theories, and see if these can be of any help in choosing one view as being better than the others, and in explaining why different, equally well informed philosophers still pick a different route on the map.[1]

§3. Here are seven things one can do when evaluating and comparing our three candidates (and metaphysical theories in general):

(i) insist on coherence/internal consistency
(ii) insist on compatibility with current science
(iii) take the default view
(iv) widen the net
(v) insist on explanatory power
(vi) insist on parsimony and simplicity
(vii) appeal to intuitions

I will now take these meta-theoretical attitudes in turn and examine each one in three steps: first, I will state the meta-theoretical criterion involved, second, I will discuss it and try to see whether it is a good one or not, and third, most relevantly, I will ask whether it allows us to select one of the three candidates as being preferable

[1]On the epistemic significance of disagreement, see Kelly (2005). See also Bennett (2008) who puts forward a meta-theoretical epistemicist view that says that (at least some) metaphysical questions have genuine objective answers but that often we cannot discover them and that consequently there is often little reason or no reason at all to go for one theory rather than the other. Compare also to Kuhn (1973): "When scientists must choose between competing theories, two men fully committed to the same list of criteria for choice may nevertheless reach different conclusions. Perhaps they interpret simplicity differently or have different convictions about the range of fields within which the consistency criterion must be met. Or perhaps they agree about these matters but differ about the relative weights to be accorded to these or to other criteria when several are deployed together." I will have more to say in Chaps. 3 and 4.

to the others. This is not a detective story so I can reveal the end of it already: there will be no winner. I have already discussed the notion of explanatory power (criterion *v*) in Chap. 4, and I will add some new considerations below. Criterion *vii* merits its own chapter and will be the focus of my interest in Chap. 6.

Note that in what follows I will focus on the role these criteria play in theory *evaluation* and especially in the process of selecting of one theory as being superior to its competitors, but we should bear in mind that as well as being criteria for evaluation of the finished products, (at least some of) these criteria are also principles that drive our theoretical choices when we are involved in the process of *building* metaphysical theories. What I say below can largely be applied to this aspect of the criteria's theoretical role as well.

§4. Insist on coherence/internal consistency. Without argument, I assume that any good candidate has to be a coherent and internally consistent view; respecting this criterion is a minimal requirement in order to be included on the map in the first place.

All three candidates of my example do accommodate this requirement. It is true that some have objected to Theory 1 that it appeals to an incoherent notion of a substratum, but this worry can be answered (see OBJ5). So, although this criterion is a good one (indeed, an obligatory one), it does not help in selecting one candidate over the others (since to *be* a candidate is to respect the criterion).

§5. Insist on compatibility with current science. According to this criterion, any metaphysical theory should at least be compatible with the most advanced physics we know. Maudlin (2007, p. 1) holds an even stronger view: "[…] metaphysics, insofar as it is concerned with the natural world, can do no better than to reflect on physics. Physical theories provide us with the best handle we have on what there is, and the philosopher's proper task is the interpretation and elucidation of these theories."

While it seems to me reasonable to require that what we do in metaphysics does not contradict what is being done in physics, there are some difficulties here, especially with the stronger version of this criterion. *Firstly*, lessons from the past teach us that the most advanced physics we know may simply be false. *Secondly*, this criterion somehow over-idealizes the state of current physics: it assumes that there *is* something like the best current physical theory and that physicists agree on what is their best theory to date (with which metaphysical theories should be compatible), but this just is not the case. *Thirdly*, it is notoriously difficult to see what metaphysical claims physical theories imply, since the equations by themselves do not say anything of relevance, and the *interpretation* of the equations is a difficult task, and one that already involves philosophical/metaphysical manoeuvres. *Fourthly*, and perhaps most importantly, as Paul (2012) remarks, the questions and problems addressed by metaphysicians are distinct from those addressed by physicists. As Paul rightly claims, there are some objects, properties, or relations that can be studied by scientific observation and measurement (microparticles, galaxies,…), but there also are many entities that are not scientifically observable in this way, like abstract objects, properties, the composition relation, the material constitution relation, the grounding relation, ersatz possible worlds, and many others. Some of these entities can manifest themselves in our ordinary, everyday experience of the world and so we can get some pre-theoretical knowledge about them via ordinary observation, but it is not the case

that we could discover further characteristics of such entities by the means of scientific investigation. "Use of an electron microscope or other measurement devices will not give us any more empirical information about the nature of composition than we can derive from everyday experience. […] There is very little about the ontological that is observable, and what is observable is observable (at least for the most part) at the level of ordinary experience. […] There isn't the faintest glimmer of an idea of what sort of instrument (much less an idea of how to build one) we could use to detect the presence of numbers, of composition, or of necessity, or whether a property is an instance of a universal or is a trope." (Paul 2012)

So, although it seems a reasonable and highly desirable thing to avoid contradictions with physics in order to gain support from it and to include metaphysics in a wider network of scientific research, the criterion seems to be a non-obligatory one, and one where we must proceed with care. Hawley (2006, p. 454) puts it in this way: "The fact that a metaphysical view seems to be part of a coherent and empirically successful scientific theory gives us good reason to think that it is empirically adequate, but […] this is very far from being a guarantee of its truth."

Here again, all three candidates seem to be on equal footing, as none of them exhibits a contradiction with current physics (especially because, as we have seen above, the questions they address are often questions that physics does not).[2]

§6. Take the default view. "If all competing views fail, take mine". This is a widely used strategy both in the literature (just think of David Lewis' "On the plurality of worlds") and in our everyday way of doing metaphysics. If all views except one present very serious drawbacks that one cannot see how to solve, then it is, quite trivially, natural to embrace the view that works best.

But, perhaps, one should only accept a view if it has real merit, rather than just by default. So, the correct way to state this criterion is to say that if the 'only' view that works is appealing regardless of the failure of its competitors, one should embrace it (Lewis would claim this to be the case as far as the plurality of worlds is concerned), but if it is only supported by the fact that its competitors fail, the right attitude to adopt is probably to refrain from holding any definite view on the issue, and perhaps only use the view at hand as a provisional hypothesis.

This criterion does not allow us to choose among our three candidates, since no two of them can be clearly eliminated.

§7. Widen the net. Do not look (only) at isolated areas on the map. Rather, when evaluating an argument or a theory, see how it fits in a more general picture. Let us consider an idealized example: if, other things being equal, one of the three candidates is compatible in a better way with neighbouring issues such as, for instance, vagueness, persistence through time, modality, or others, then it is a good reason to choose it over the other two, since it provides us with the possibility of having not only a good 'local' theory but also a coherent 'global' metaphysical/philosophical picture.

[2]Other metaphysical debates might be perhaps more easily influenced by physics, like the presentism/eternalism controversy. But even there, I think that the arguments from physics are very tricky ones, and not decisive for the three reasons given above.

An interesting question is whether, in the last sentence, I should have written only "metaphysical" rather than "metaphysical/philosophical". What I have in mind is the case of arguments such as the one in Mark Heller's "The Immorality of Modal Realism, or: how I learned to stop worrying and let the children drown" (Heller 2003). (The case of this argument is interesting not only for its own sake and as an example of a general point about widening the net, but also because it is directly relevant with respect to the case of our three candidate theories, since two of them seem to be committed to modal realism *à la* Lewis.) In this paper, Heller intends to show that Lewisian modal realism has consequences that are morally unacceptable. Suppose we accept Heller's arguments and his conclusion. One could ask: isn't Heller widening the net *too much*? Is it relevant, when evaluating a metaphysical theory, to consider what moral consequences it has? To take a parallel case in physics, if it turned out that the best interpretation of quantum mechanics is the Many-Worlds interpretation (see Everett 1957; De Witt and Graham 1973), would we reject it or consider it as being a worse theory because it has some immoral consequences? The general problem here is: how far should we widen the net?

One thing strikes me as quite trivial: the 'widen the net' criterion *is* a good one. It *is* a virtue of an argument or a theory that it can be fruitfully included in a wider network of human knowledge; indeed, this parallels the requirement for internal consistency, except that it is internal consistency of something like a 'theory of everything' that we are aiming at here.

But again, how wide should we go? Heller definitely believes that moral issues are relevant, and even *decisive*, for metaphysical theories[3]; his claim could not have been clearer or stronger:

> I will argue that if modal realism is true, there are at least some cases in which it is permissible to let drowning children drown when it would be easy to save them. But this is not permissible. By *modus tollens*, then, modal realism is false. (Heller 2003, p. 3)

Besides, Heller also claims[4] that the 'parallel' case from physics is *not* parallel. If physics had the same moral consequences modal realism has, this fact would *not* count as an argument against it.

I think one has to be careful about the implications of the moral implications of a metaphysical theory. Granted, if any theory entails any falsehood, then a *modus tollens* can be drawn. But clearly, in the case of Heller's argument and most likely in all similar cases, one does not get *any* moral consequences from modal realism *alone*, rather one gets moral consequences from modal realism *plus* some additional moral thesis or presuppositions. In Heller's case, these include at least:

- it is not permissible to let children drown when it would be easy to save them (Heller 2003, p. 3)

[3]See also Adams (1974), and Beedle (1996). Buddhist metaphysical claims also seem to be mainly driven by Buddhist ethics.

[4]Heller (2003, Footnote 10).

- mere spatial proximity should not have any moral significance; applied to the case of Lewisian modal realism: being located in the same world should not have any moral significance (Heller 2003, p. 4)

It is only if these (very sensible, I don't dispute that) moral thesis are accepted that Heller can make his point against modal realism (Heller agrees with that, of course). And, in general, in any place where we consider consequences that a metaphysical theory has with respect to other claims or theories (vagueness, personal identity, modality, …), this will always involve the acceptance or rejection of some additional thesis external to the theory being considered. This is why, while widening the net is a very virtuous attitude when evaluating metaphysical theories, the consequences of such 'widenings' are to be taken, I think, as less directly relevant than direct consequences of the theory alone, since the former involve additional thesis that may always, in principle, be considered as controversial. This, in itself, does not of course make this meta-theoretical criterion any less interesting, it just makes it harder to work with. What makes it even harder, especially if non-metaphysical issues are to be taken into account, is the pragmatic difficulty of determining all the connexions of one theory to all the other issues, arguments, and theories. One person could certainly not do it, but perhaps our collective efforts can.

Let us now see whether and how this can help us with our three candidates. One example: two of them are strongly connected to a certain theory of possible worlds in order to be able to face some objections. "Bad point for them", says the opponent. "Good point for this theory of possible worlds", says the defender. To arbitrate this dispute, we would need to have something like "the whole picture"—remember, at the very beginning, I suggested that in order to be able to evaluate the three candidates they have to be 'complete enough', that is, not only theories that answer the question of the nature of properties, but that also answer the question of the nature of objects, for theories whose scope is too narrow are hard to evaluate for precisely the reason that we would constantly need to look elsewhere to see how they fit with neighbouring issues. This alone shows that the "widen the net" criterion is a good one, and it shows where it really leads: it asks us, in the end, to compare not just our three candidates but to compare and evaluate only 'theories of everything'—that is, only fully global pictures. This a mostly virtuous aim, and I hope that, one day, we'll be able to do it. In the meantime, we can only consider partially global views, and there does not seem to be anything decisive in favour of one of our three candidates here.

§8. Insist on parsimony and simplicity. Many metaphysicians appeal to and use the principle of simplicity and parsimony in their work, but few of them explicitly say why. "If you can do it with less, don't do it with more", the slogan goes (often only assumed in silence). But why? And what does "less" and "more" mean? Several criteria can be distinguished here:

(i) simplicity of the structure of the theory
(ii) number and complexity of primitives in the theory
(iii) qualitative parsimony (concerns the number of kinds of entities)
(iv) quantitative parsimony (concerns the number of entities)

These four different criteria are, of course, related. For instance, with greater complexity with respect to (i) it is likely that a theory will be able to be more parsimonious with respect to (ii), and vice versa. The same may apply for (iii) and (iv) as well: more kinds of entities, as in the case of Theory 1 that includes particulars *and* universals, perhaps allows for being more parsimonious with respect to the number of these entities than, for example, Theory 2 that includes only particulars—tropes—where for each instantiation of a universal (one entity) there has to be a different trope (i.e. there are as many entities as there are instantiations of the one universal).[5] Thus, being more parsimonious with respect to some of the criteria above is very likely to force one to be less parsimonious with respect to others. So, the question is, are some of these criteria more important than others? Many think so as, for instance, does David Lewis:

> I subscribe to the general view that qualitative parsimony is good in a philosophical or empirical hypothesis; but I recognize no presumption whatever in favour of quantitative parsimony. (Lewis 1973, p. 87)

Nominalists would probably generally agree (see also for instance Rodriguez-Pereyra 2002, p. 202). Here is a naïve but genuine question: are nominalists nominalists because they hold that qualitative parsimony is more important than quantitative parsimony, or do they hold this view because they are nominalists? It should be (and probably is) the former, otherwise their overall strategy could be regarded as question-begging. But then, if this meta-theoretical criterion is not supported by their ontology, by what *is* it supported? Note that the price to pay for this criterion is not only that their ontologies are presumably less quantitatively parsimonious, but also that they usually have a more complex structure (compare, for instance, the structure of Rodriguez-Pereyra's resemblance nominalism, which includes a hierarchy of pairs of resembling particulars (Rodriguez-Pereyra 2002, Chap. 9)) in order to account for one particular's being F, to the structurally simpler view that the particular instantiates a universal). The support for this claim comes, as Rodriguez-Pereyra (2002, p. 202) puts it, from the view that "[…] the *generality* of metaphysical theories makes the existence of a certain *kind* of entities matter more in assessing their truth or falsity than the number of entities of that kind."

Not all nominalists agree. For instance, Melia (2005) prefers to take the route of parsimony with respect to the structure of his nominalist theory, but completely sacrifices parsimony with respect to the kinds of entities included in the theory—indeed, according to his view, there are (almost?) as many kinds as there are individuals. The very general idea behind this sort of meta-theoretical attitude is that qualitative parsimony *is* a virtue of a theory and we should be as parsimonious as possible, *but no more*—that is, if introducing a new kind of entity is justified by its usefulness in generating a structurally simpler view, then it *should* be introduced. Thus, ironically enough, Melia's nominalism joins the side of the friends of universals—for it is precisely this kind of considerations that defenders of universals

[5]Note, importantly, that this makes Theory 1 more quantitatively parsimonious only if there are *finitely* many universals in the actual world. (see Rodriguez-Pereyra 2002, p. 204–).

use in order to support their view: universals are not superfluous entities that can be dispensed with since they do useful work in allowing for a powerful and structurally simple theory (here (i) and (iv) are privileged).[6]

Even if there is disagreement as to what sort of simplicity and parsimony should be privileged, everybody seems to agree that it is a general virtue of a theory to exhibit some sort of it.[7] Why is simplicity and parsimony interesting? Clearly, there is a pragmatic reason: a simpler theory will be easier to work with, to manipulate, to understand, to explain to others, and so on. While this certainly is an interesting feature of a metaphysical theory, it can hardly be a reason to choose one as being better than the others—such a pragmatic advantage is more of a welcome by-product of simplicity and parsimony than the main reason for desiring it. So, perhaps the criterion of simplicity and parsimony is desirable because it is truth-conducive? Could it be that a simpler theory is a better guide to metaphysical truth than a more complex one? But why? What is it that makes a simpler and more parsimonious theory more likely to be *true* than a more complex one? Perhaps behind this claim lies the general idea that the world *is* as simple as possible—as Newton put it:

> We are to admit no more causes of natural things than such as are both true and sufficient to explain their appearances. To this purpose the philosophers say that Nature does nothing in vain, and more is in vain when less will serve; for Nature is pleased with simplicity and affects not the pomp of superfluous causes (Newton 1687, rule I)

The trouble with such a view is that it is under-motivated. Why would 'Nature prefer simplicity to complexity'? What guarantees that metaphysical truths are simple rather than complex? Isn't it more plausible that to assume such a claim does no more than to express one's own (and not Nature's) preference for simplicity and parsimony? I think it is, and for good reasons: the pragmatic one mentioned above, and the fact that many of us *just like it*. So, let us sum up where we are: first, the criterion of simplicity and parsimony *is* taken as important by virtually all metaphysicians but, second, there is no consensus as to what *kind* of simplicity and parsimony is the one to be preferred, and third, the claim that this criterion is truth-conducive is under-motivated.[8] What I want to suggest is that it is, in a good sense, 'wishful

---

[6]Tim Maudlin puts the point as follows: "It is not clear that [Ockham's Razor] can withstand much critical scrutiny. If by 'necessitas' [in "Entia non sunt multiplicanda praeter necessitatem"] one means *logical* necessity, then the Razor will land us in solipsism. But if one means something milder—entities ought not to be multiplied *without good reason*—then the principle becomes a harmless bromide: nor should one's ontology be *reduced* without good reason." (Maudlin 2007, p. 3)

[7]In conversation, John Bigelow expressed the view that the best theory would be one that embraces *both* tropes and universals. While this can sound very unparsimonious and unappealing at first sight, it is no more than following a little further the idea that explanatory power and simplicity of structure of this 'mixed' view is more important than qualitative parsimony.

[8]Sober (1981) and Rodriguez-Pereyra (2002, p. 207–) discuss the idea that a (qualitatively) more parsimonious ontology is more likely to be true than a less parsimonious one because a conjunction (of claims) always has a lower probability than any of its conjuncts (if these are mutually independent). But this only works, following what Rodriguez-Pereyra rightly says, if it is possible to attribute comparable probability values to the axioms of a theory—but it seems to me that

thinking': the requirement for parsimony and simplicity comes from *us* rather than from the metaphysical reality.

§9. Insist on explanatory power. Of course, this is a good meta-theoretical criterion. After all, that's what we have a metaphysical theory *for*. What we do in metaphysics is to build theories that explain or analyse what we want them to explain or analyse, and if they don't do that, they can hardly even count as candidates. Thus, many of us often insist on explanatory power as being a very important meta-theoretical criterion for choosing one view rather than another—while acknowledging that explanatory power isn't everything: for instance, there are many who agree that Lewisian modal realism *does* greatly satisfy this criterion, but nevertheless very few are ready to embrace a plurality of worlds.

All three candidates do satisfy this criterion, all three have the explanatory power to face the questions we asked, all three deliver the goods. An interesting question is *how* they gain this power, and the answer lies in the second general remark made above in §2:

- The *primitives* of each candidate do a big part of the job, when it comes to answering the initial question(s) and replying to objections. It is only *if* we grant the views their primitives (and ontological commitments) that they can face all the objections.

Indeed, without its primitives, none of the three candidates could do the job it does, and, of course, this is perfectly normal, since that is why their primitives are introduced in the first place; if they weren't doing an important job in the explanation provided by the theory, there would be little justification for introducing them. This is true here, as it was true in the case of other theories from Part I. Comparing the explanatory power of good candidates (that is, candidates that do give a complete answer to the question(s), and so cannot be distinguished on this base) amounts, then, to comparing their primitives. But how to do this? (As we have seen in Part I, one thing that makes the comparison of primitives hard is that, in many relevant cases, the primitive of one of the candidates is functionally equivalent to a primitive of another.) Often, the 'explanatory power' criterion is of course a good one, but it suffers from the same problem as the criterion of 'internal consistency'. The good candidates all have it, and are thus distinguished from bad ones that do not deliver the goods we asked for. Once this selection is done, comparing the explanatory power of the good candidates amounts mostly to comparing their primitives.

To have a clear example in mind, take a red apple that is the same shade of red as my neighbour Cyrano's Ferrari. Both objects 'have the same property'—a phenomenon that is analyzed differently by our three theories. According to the theory of universals, objects like apples and lizards are substrata that instantiate

(Footnote 8 continued)

asking, for instance, whether the existence of tropes is more or less *probable* than the existence of universals is wrongheaded, for how could any such claim ever be justified?

spatio-temporal, multiply locatable (repeatable) universals, where two objects 'share the same property' iff they both instantiate the (numerically) same universal. According to trope theory, there are no substrata, objects are 'just' bundles of compresent non-repeatable, non-multiply locatable, properties (tropes), where two objects 'share the same property' iff they both have among their members/constituents/parts numerically different tropes that are exactly similar. According to resemblance nominalism, two objects 'share the same property' iff they are both members of the same resemblance class.

As different as these two answers are, they have something in common: they all solve the problem by postulating a primitive solution to it—they use *problem-solvers*, a tool now familiar to us from the discussions in Part I, except that now the theories under scrutiny are clearly *not* equivalent. The *relation of exact resemblance* between tropes, the *fact that two objects resemble each other*, as well as the *instantiation of the same universal* are primitives that each theory appeals to, and these primitive theoretical tools are precisely what allows them to provide an answer to the problem of attribute agreement. When one asks "In virtue of what does the apple and the Ferrari have the same property of being red?", Theory 1's answer is "In virtue of instantiating the same universal of redness". One can then ask further "In virtue of what are two instances of redness instances of redness (the very same universal)?" *That*'s a primitive. Theory 2's answer to the initial question is "In virtue of containing exactly similar tropes of red". And, again, one can ask further "In virtue of what are two exactly similar tropes of red exactly similar?" *That*'s a primitive. Similarly for Theory 3, whose initial answer is: "In virtue of being both members of the same resemblance class". Question: "In virtue of what are two objects members of the same resemblance class?" Answer: "In virtue of the fact that they resemble each other." Question: "In virtue of what do they resemble each other?". *That*'s a primitive.

All three theories thus answer the problem of attribute agreement in a primitivist way. Ultimately, the problem is "solved" by postulating a tool, and by postulating that the tool is primitive and not further analysable in the framework in which it is postulated. (Note, for instance, that while *exact resemblance* is thus a primitive under the trope-theoretic framework, it is not primitive under the universals view where exact resemblance can be analysed in terms of instantiation of the same universal(s). Each theory has *its* primitives, and what one takes as primitive the other typically does not.)

Here is another example. The friend of the universals view is often charged with objections concerning the relation of instantiation. If the relation of instantiation relates the substratum to its properties, what relates the substratum to the relation of instantiation? A threat of regress and related worries arise. It is the nature of the relation of instantiation that is in question, which is no surprise: after all, it is a theoretical tool, concerning which we do not have any common sense intuitions (mostly because we have no common sense intuitions at all on such matters, as we will see in the next chapter). The relation is one that somehow connects universals and substrata, but not in a way normal relations relate—which makes objectors raise their eyebrows. The strategy of the defender of the universals view is simply to

meet such objections head-on. Yes, instantiation relates but it is not a relation, it is a "non-relational tie" (Strawson 1959), it is an ontological glue that makes properties 'stick' to substrata (Broad 1933), and it is 'self-adhesive' in the sense that it does not require a further relation to relate it to the substratum (Armstrong 1978). Such a reply to the objector's worries has two features that I am interested in. First, it is a primitivist reply—instantiation is a primitive that is postulated to have such and such theoretical virtues, which allow the friend of universals to answer many of her theoretical challenges (see also the sketchy list of objections in §2 above). Thus, here again, the answer to the problem is, at the end of the day, a primitivist one. Secondly, and very importantly, this way of dealing with a problem is perfectly acceptable, and is common methodology in metaphysics. If a metaphysician wishes to introduce a primitive tool in her theory because she needs it to solve a number of problems then it is, of course, perfectly all right that *she*—and *not* her opponents—is allowed to specify the nature and the features of *her* primitive. A primitive like the relation of instantiation is *not* like a rock made of gold that one finds in a mine and whose features one can examine to see how rich one just got. Rather, such a primitive is a theoretical postulate, whose features are postulated by the metaphysician who introduced the primitive in the first place—consequently, she can postulate it to have all the features she wants it to have. And of course, she should allow her opponent to do the same with respect to *her* primitives.

The trope theorist surely does behave in the same way. One traditional objection to the tropes-bundle-theoretic view is that it faces a version of the 'problem of naturalness' sometimes also called the 'imperfect community problem'. Here is a way Manley (2002, p. 84) puts the objection: "Consider the resemblance class of tropes (in the actual world) that has as members all and only the pink color tropes, the baby-blue color tropes, and the purple color tropes. These all resemble each other to a good degree and no non-member resembles each of them to that degree. But this is no property class worthy of the name 'natural'". Somehow, we need to distinguish between 'natural' properties, like the property of *being pink*, and other 'non-natural' properties like the property of *being either pink or baby-blue or purple*. But, the objector points out, if one uses resemblance classes to play the role of properties in the trope theorist's ontology, one cannot make this distinction. The reason behind this flaw of trope theory is that resemblance is too vague and that it holds not only between two tropes of the very same shade of pink but also between a trope of pink and a trope of purple as well. But there is an easy solution to the problem, as Manley points out: the worry does not even arise if we use *exact resemblance* rather than just *resemblance*. Exact resemblance is, indeed, what the objector asks for when raising her objection, since *exact resemblance classes* will only, by definition, give rise to perfectly natural properties. It is, then, enough for the trope theorist to give us exact resemblance rather than resemblance, and it is something that is very easy to give: resemblance, and exact resemblance, are both primitives postulated by the metaphysician who needs them to make her theory work, and if she claims that there is a primitive such as exact resemblance—one that gives rise to precisely the resemblance classes she needs—the objector's worries do not even arise: problem solved! It is thus in a primitive way that the theory faces its

theoretical challenge. This may not satisfy the objector, but as in the case of the theory of universals, from the methodological point of view we can ask: why wouldn't the tropes-bundle theorist have the right to choose what her primitives are and what they can or cannot do? After all, they are *her* primitives, *her* tools that she decides to use in order to make her theory work, and since they are not golden rocks but rather theoretical postulates, there is no reason why she couldn't specify how they are and what they do.

§10. We have seen that, in several crucial places, our three theories answer their theoretical challenges in a primitivist way, similarly to the cases of theories we have seen in Part I. If you have a look at the list of objections and replies in the tables in §2 above, you'll quickly realize that I have *not* carefully chosen my examples to fit my overall claim—rather, such a phenomenon is largely manifest in the whole debate between the three views. Without their primitives, theories like trope theory, resemblance nominalism, or the theory of universals could not even begin to work, and could clearly not face their theoretical challenges and objections (again, as we have already seen in the case of many other theories in Part I). As I have already mentioned, I think that this is not at all a surprising or objectionable claim. Indeed, not only it is a common methodology to introduce primitives that do heavy-duty jobs, but it even seems that it is the only reason to postulate them, for otherwise there would be little justification to bother with them in the first place. If it weren't for some important theoretical job to be done, we would clearly feel no need to postulate the existence of an entity such as a non-relational instantiation tie. Primitives are acceptable in our theories precisely because they do an important job.

What to do, then? In this chapter, I have raised worries with respect to meta-theoretical criteria for theory evaluation such as a theory's simplicity and parsimony. These remarks were not intended as fully developed and conclusive discussions of these important methodological and meta-theoretical criteria, rather, their role in what I wanted to achieve in this chapter is to strengthen the main general point I want to make. Perhaps my main point is actually not very controversial and perhaps not even very original but it is hardly ever acknowledged and seriously taken into consideration in the debate between trope theorists, nominalists, and friends of universals (as well as relationists, substantivalists, perdurantists, and endurantists): if one grants one's opponents the same right to introduce primitives as one grants oneself, all theories will quite nicely end up being able to face all of their objections and theoretical challenges, in a primitivist way. Granted, the three views used as examples in this chapter—contrary to the theories from Part I—are clearly different, and there is no metaphysical equivalence claim in the neighbourhood of what I am claiming here, but there still seem to be no strong ground to choose one theory as being superior to the others. Behind this claim lies the idea that explanatory power is by far the most relevant and useful meta-theoretical criterion for theory evaluation, and consequently, that if two views get their power from their primitives, they end up being on a par at the end of the evaluative process—after all, metaphysical theories are there to do a job, and if they both do their job in a primitivist way, one can hardly be said to have more explanatory power than the other. As a result, there is no good reason for making a choice.

Indeed, the primitives involved in our metaphysical theories are problem-solvers: primitives that are there to solve a problem (like the relation of exact resemblance between tropes, or the instantiation of the same universal). When facing their theoretical challenges, our theories introduce a primitive with the same function: primitively answer the question "In virtue of what does the apple and the Ferrari have the same property of being red?" Problem solved. With a problem-solver. Primitives are thus absolutely crucial to the success metaphysical theories can have when trying to solve philosophical problems. But not only this: primitives are also 'points of contact' between theories. Think about the theories discussed in this chapter and the problem of attribute agreement. The primitives, while being clearly different, have the same overall function within the theory: they explain how two objects can share the same property. In a very general sense (*en passant*, too general to justify any claims of equivalence), our three views thus answer the question in the same way (namely, in a primitive way). This is where sceptical worries can start to arise. Indeed, if both theories provide a solution to a problem in a primitivist way, what difference does it really make, with respect to the problem at hand, to choose one theory rather than the other? If, at the end of the day, rival theories like trope theory or the theory of universals both solve the philosophical problem we want them to solve by using their primitives, why should we prefer one over the other?[9]

As announced, the purpose of this chapter was mostly critical and negative. Strengthening the claims concerning equivalence, primitiveness, and explanatory power from Part I, we have seen that even in cases of clearly non-equivalent theories, these considerations lead us to an impasse: we are not able to choose one theory over the others, at least not if we grant all sides of the debate the same right to introduce primitive problem-solvers. What to do? How should we chose between different—non-equivalent—metaphysical theories? In Chap. 3, I will try to provide an answer, and defend a positive claim. But before I do so, I need to discuss the second most important criterion for theory choice and theory evaluation (after explanatory power): compatibility of our metaphysical theories with philosophical or pre-philosophical intuitions. This is what the next chapter is devoted to.

---

[9]This worry about theory choice is to be distinguished from a concern about an interesting and strong relationship there is between my way of establishing claims of metaphysical equivalence and the attribution of theoretical virtues, such as those examined in this chapter, to theories. Indeed, there is a 'mapping' between equivalence in my sense and the having of such theoretical virtues: if two theories are equivalent, their theoretical virtues will most often be the same, and where they are not equivalent, the virtues will be different.

This does *not* mean, however, that an account of metaphysical equivalence based on the attribution of theoretical virtues would turn out to be equivalent to my account of equivalence, for even if there is such a relationship between the two types of claims, they are clearly different: examining the theories' primitives and their structure, and looking at the way the theories do their job and face various arguments/objections, etc. is not the same thing as evaluating what their theoretical virtues such as those discussed in this chapter are. Furthermore, the latter method of trying to establish equivalence claims would not be a very good one, since, as we have seen, the evaluation, attribution, and comparison of such theoretical virtues is a very tricky affair.

Chapter 6
On Intuitions and Experience—The Marriage (and Divorce) of Phenomenology and Metaphysics

Abstract The link between intuitions and metaphysics is a strong and important one, and there is hardly any metaphysical discussion where intuitions do not play a crucial role. In this chapter, I first offer some quite general critical considerations about the role of intuitions in metaphysical debates, and I then focus on a particular kind of intuitions, namely those that come, at least partly, from experience. There seems to be a route from experience to metaphysics, and this is the core of my interest here. In order to better understand such 'arguments from experience' and the kind of relationship there is between this type of intuitions and metaphysical theories, I examine several cases where a kind of experience-based intuition seems to motivate or support a metaphysical theory. At the end of the day, I argue that this route is a treacherous one, and that in all of the cases I concentrate on, phenomenological considerations are in fact orthogonal to the allegedly 'corresponding' metaphysical claims.

§1. More often than not, metaphysical theories are counter-intuitive. Some tell us that there are weird material macroscopic objects, like an object made up of the top half of the Eiffel Tower, the north face of Everest, the last living polar bear, and Socrates' nose. Other tell us that time does not pass and that ordinary objects—including ourselves—do not persist through time, or for that matter, that time does not exist and neither do ordinary objects, including ourselves. Some claim that there is no universe around us at all and that the world we seem to be living in is actually akin to a dream. Yet other tell us not only that our universe, time, us, and weird macroscopic objects *do* exist, but that so do infinitely many other counterpart universes, as real as our own.

As a metaphysician, facing such strong and striking claims, one may wish to react in various ways. One may want to reject these claims (or, at least, incredulously stare at them). More daringly, one may want to reject the intuitions that these claims violate. Or, one may try to show that these claims actually do *not* violate any (real and good) intuitions. Alternatively, one may try to show that one's opponent's theories are as counter-intuitive as one's own. Sometimes, things go the other way around: one *starts* from a (real and good) intuition, and develops a metaphysical theory based on it.

One way or another, the link between intuitions and metaphysics is a strong and important one, and there is hardly any metaphysical discussion where intuitions do not

J. Benovsky, *Meta-metaphysics*, Synthese Library 374,
DOI 10.1007/978-3-319-25334-3_6

play a crucial role. In this chapter, I will first offer some quite general critical considerations about the role of intuitions in metaphysical debates, and then I will focus on a particular kind of intuitions, namely those that come, at least partly, from experience. (Actually, it seems to me that most, if not all, of our intuitions are like this, but I shall not argue for *this* claim here.) There seems to be a route from experience to metaphysics, and this is the core of my interest here. In order to better understand such 'arguments from experience' and the kind of relationship there is between this type of intuitions and metaphysical theories, I shall examine several cases where a kind of experience-based intuition seems to motivate or support a metaphysical theory. At the end of the day, I shall argue that this route is a treacherous one, and that in all of the cases I shall concentrate on, phenomenological considerations are in fact orthogonal to the allegedly 'corresponding' metaphysical claims.

We will thus see not only that the various meta-theoretical evaluative criteria from Chap. 1 do not allow us to make a choice between the competing theories I used as my examples, but also that a theory's compatibility with philosophical or pre-philosophical intuitions fails to fulfil that role as well. At the end of the day, as a cumulative effect of the reasons and arguments offered in Chaps 1 and 2, we will then find ourselves in the uncomfortable position of not knowing how to select one theory over its competitors—which is exactly the point where I want to get to before embarking on a new line of reasoning in Chap. 3.

§2. In the previous chapters, I insisted on the crucial role played in metaphysical theories by primitives (especially, primitive problem-solvers). Often, one finds a primitive counter-intuitive. What does it mean, and how is it relevant? Rodriguez-Pereyra, while defending resemblance nominalism, writes:

> I cannot see why theories in general should preserve intuitions, that is, pre-theoretical and uncritical beliefs. No doubt there are areas, like some areas of Philosophy of Language, where intuitions are of paramount importance. Do definite descriptions *name* things? Are proper names rigid designators? Is a 'simple' sentence like 'Superman went into the telephone booth' extensional? Intuitions are of great importance in answering these and similar questions having to do with meaning. The reason for the importance of intuitions in this area is that, after all, meaning is something *we* do and so we can reasonably expect that our intuitions about meaning will be approximately correct.
>
> But with metaphysical theories about the basic structure of the world, like Resemblance Nominalism and other solutions to the Problem of Universals, there is no reason to expect that our pre-theoretical beliefs and opinions will be true. […] [O]ne should always keep a critical eye upon intuitions and be ready to discard those that are not validated by a rational and critical assessment or those that conflict with scientific or philosophical theories. Merely preserving certain intuitions does not make a theory better. (Rodriguez-Pereyra 2002, p. 217)

Rodriguez-Pereyra describes intuitions as being "pre-theoretical" and "uncritical"—he takes them to be opinions/beliefs of untrained common sense. With respect to this sort of intuitions, what he says is largely right. Such intuitions do sometimes count, but not in all areas of philosophy, and only to some extent. And basic metaphysics probably is among the areas of philosophy where such intuitions count the least. Furthermore, even if these intuitions were to count, their role would not be trouble-free. They are not universally shared, and are even variable over time for one thinker—consequently, it is hard to use them to establish any permanent and

universal result. Also, it is notoriously difficult to distinguish between 'good' and 'bad' intuitions—we often have conflicting intuitions, or weaker and stronger intuitions, and it is far from clear how to sort them out as being more or less fundamental and/or reliable. Thus, relying on such intuitions when doing metaphysics is really relying on something rather unreliable.

But there is also another kind of intuition, that is perhaps more serious[1]: the intuitions of the philosopher who considers a thesis, proposition, concept,… and looks for an intuitive understanding of it. Here, the idea is that, mainly by considering particular cases, examples, or by making thought-experiments, one can make more precise and more salient some intuitions that are stronger than mere uncritical opinions of common sense, and that arise from careful consideration of the case or thought-experiment. For instance, imaginary cases or Star Trek stories of duplication of persons in the debate about personal identity allow us to give rise to some more carefully formulated and useful intuitions that can probably do some helpful work in the understating of our concept of a person and her conditions of persistence through time.

But even here, similar worries apply. Firstly, as before, even these more serious intuitions do not seem to be relevant in the field of basic metaphysics, which is just too abstract and theoretical for any useful intuitions to arise. Of course, basic metaphysical theories may have implications with respect to other theories, like theories of persistence through time of persons, and *there* useful intuitions can arise—so, it is by *widening the net* (see Chap. 5. §7) that one can include intuitions in the game. But, secondly, even intuitions of this more serious kind suffer from being too unsettled and variable from one thinker to another and over time for one and the same thinker,[2] and even here conflicting intuitions, good and bad intuitions, as well as weak and strong intuitions can arise—thus, for similar reasons as before, they do not seem to be a very reliable guide, even if they are more reliable and at least to some extent more useful than the uncritical and pre-theoretical intuitions of the first kind.

My concern here is with intuitions in the field of metaphysics, but similar claims have been raised in other philosophical domains as well, for different reasons. For instance, Machery, Mallon, Nichols, and Stich (2004) offer persuasive evidence that *semantic* intuitions are culture-relative, thus making it difficult to rely on such intuitions when developing a philosophical theory of reference. Also, when it comes to *thought-experiments* and intuitions involved in these thought-experiments, Swain, Alexander, and Weinberg (2008) show that *depending on the order* in which subjects are presented with thought-experiments, their responses vary—thus, intuitions are shown to be easily manipulated (the intuitions involved in this case were *epistemic* intuitions, arising from Lehrer (1990, pp. 163–164) Truetemp Case). As they put it (§1): "Intuitions track more than just the philosophically-relevant content of the thought-experiments; they track factors that are irrelevant to the issues thought-experiments attempt to address. […] Such sensitivity to irrelevant

[1]The point that useful intuitions are those of a 'more serious' kind than ordinary intuitive beliefs is also, in a different way, discussed by Bealer (1987, 1988).

[2]See Machery (2004), Swain (2008), Weinberg (2010).

factors [socio-economic status, cultural background, or the order in which one considers various thought-experiments] undermines intuitions' status as evidence. Evidence so unstable risks being discounted as not being truly evidence at all. Furthermore, given that intuitions vary in these ways, there is unlikely to be a fixed set of intuitions about a particular thought-experiment to which we can appeal."

A possible reaction a defender of the use of intuitions in philosophy could have would be to insist that *philosophers*, since they are experts, have better intuitions than any pre-theoretical intuitions of people who are, say, presented for the first time with a thought-experiment. The idea here is not to count as an intuition the first thing that springs into our minds when we are first presented with such cases, but rather to think about them and discover a kind of 'conceptual intuition'—a philosopher's, that is, an expert's intuition.

But it is far from being clear that being experts gives us, *qua* philosophers, any advantage. One way to raise this critical point is mentioned by Scholl (2007, p. 585) who focuses on the use of thought-experiments involving imaginary examples such as, precisely, cases of fission of objects or people. He says: "This analysis might seem to lead us to question the use of such outlandish examples in cases involving object persistence. […] The outlandishness per se may not be problematic, but these scenarios may violate assumptions about the world that are made in a reflexive way by our perceptual and cognitive processes. This is consistent with previous discussions of the danger of such thought experiments. […] Such perceptual illusions may, in the service of metaphysical theorizing, become cognitive illusions that can lead us astray." In short, the outlandishness of such scenarios may introduce even more possible bias than more mundane experiments. Weinberg et al. (2010) object to the claim that philosophers, as experts, have intuitions any better than the layman's for a similar reason: expertise, in general, is not a cure for cultural or other biases, and there is no reason to think that philosophers are an exception. A higher level of expertise, as they argue, can actually be a source of more ways to go wrong, when appealing to one's expert intuitions. Suppose, for example, that a philosopher has at her disposal better conceptual schemata than a non-expert, thanks to her training. While such schemata might indeed help to avoid some of the misleading folk intuitions, working with such schemata and concepts is working with heavy cognitive machinery that might bring new possibilities of biases on its own—one may be introducing a new source of error without realizing it.[3] Besides, there is no reason to think—and even less reason to think there is a proof—that philosophers are immune to the defects of other scientists' or non-experts' thinking, like overgeneralization, overconfidence, belief bias, or belief perseverance. While believing

[3]Compare to Walton (2008, p. 110), commenting on the alleged counter intuitiveness of his 'transparency claim' about photographic representation: "So shall we argue about whether [the transparency claim] *really is* counterintuitive? It would be better simply to recognize that intuitions are largely reflections of one's currently internalized theoretical commitments (there being no such thing as entirely *pretheoretical* intuitions), and that whatever authority one accords them amounts to resistance to theory change simply because it is theory change."

that we, as philosophers, have finely tuned deep intuitions that are better than the layman's, we might actually only be reformulating early intuitive impressions.[4]

What about intuitions and our three candidate theories from the previous chapter? Rodriguez-Pereyra, while defending his version of resemblance nominalism, thinks that his own view suffers from being less intuitive than its competitors (but he does not think that it is a strong drawback, following the considerations above). Here, I do not agree. Rodriguez-Pereyra's main reasons for this are that resemblance nominalism is committed to modal realism, while our intuitions are actualist, and that his view does not follow our intuitions to the place where the having of a property is an intrinsic matter. This, I think, is correct—at least to the extent to which any intuitions concerning such matters can be relevant. But the bundle theory with tropes requires modal realism as well (see Manley 2002), and the existence of a substratum, or of multiply locatable spatio-temporal entities such as immanent universals, certainly have a lot of incredulous stares of their own to face. Every one of the three candidates we have seen has its primitives and ontological commitments, and every single one of them has been, at some point, labelled as counter-intuitive. I think that this is not surprising: a primitive or unexplained and unanalysed claim is easily found counter-intuitive since, by being primitive, it sounds mysterious (at least to the ones who don't find the primitive intuitively intelligible). Thus, generalizing perhaps just a little too much, primitives are often counter-intuitive in virtue of the simple fact that they are primitives. Consequently, and unsurprisingly, every theory bears its amount of counter-intuitiveness, and our three theories can hardly be distinguished on this ground.

Perhaps some will think that some counter-intuitive claims are *more* counter-intuitive than others. One might think, like Rodriguez-Pereyra (2002, p. 202) seems to think, that modal realism is the most counter-intuitive claim in the neighbourhood. But one might also think that an unknowable substratum is worse. Or one might find counter-intuitive the idea that there exists nothing more than properties, bundled together, and that objects are 'made out' of properties. I could tell you now which one of these claims I find the most counter-intuitive, and you could do the same, and it is very likely that we would not agree. What then? How can we use intuitions as good criteria for evaluating metaphysical theories? It seems we cannot.

§3. The considerations I raised in the preceding section are quite general, and even though I think they are correct, they probably are too general to be sufficient in order to establish a conclusive, critical claim about the role of intuitions in metaphysics. In what follows, I will focus on more detailed and particular cases of a type of intuition that play an important role in many debates in metaphysics: cases where one could think that there is an argument in favour of a metaphysical theory coming from experience.[5] I will argue that all these alleged arguments fail.

[4]Weinberg et al. (2010) offer a lot more detail (and more criticism) on these and related issues.

[5]In these examples, I will simply assume that hard-core scepticism about the existence of the external world is false. I will assume that there is a world existing independently of us, and that, in some way or another, we have experiences of it.

Let us consider, as a first quick example, the debate between the friend of unrestricted mereological composition, and the one who thinks that composition is restricted. Suppose that the friend of *un*restricted composition says: "I see an apple on a table. I see an apple, and I see a table. But there is also this one object made out of the top half of the apple and the left half of the table. I see this object as well, it's right there, in front of me. Let me call this object 'Bernard' and let it be my favourite object of the day." What the defender of unrestricted mereological composition (who'd embrace this type of argument) wants to say here is that *we see* Bernard, it's there, just in front of us. We can touch it, and smell it as well. So, there is no reason for saying that only the apple and the table exist, while Bernard does not. We experience all three objects equally well, so we should treat them accordingly as ontologically equal. Perhaps this is a simplified picture of what a friend of this line of argument could say, and of course not everybody would agree on the datum—namely, that we see Bernard equally well as the apple and the table. Indeed, this line of argument can be reversed and can be taken to purportedly show the exactly opposite claim—that since we do *not* perceive Bernard in the same way we do perceive the other two objects, we should *not* put it ontologically on a par with them (that's something a friend of restricted composition could claim).

But, as Merricks (2001, p. 9) rightly remarks, these considerations about our perceptions of Bernard, if taken as leading to the doctrine of unrestricted (or restricted) mereological composition (depending on what you think of the perceptual datum), is entirely misguided: our visual, olfactory, and tactile evidence would be exactly the same whether Bernard existed or not. That is, whether the top half of the apple and the left half of the table compose an object is an ontological question, and a mereological claim that is entirely neutral with respect to what we see. Our visual phenomenal experience when contemplating this scene would be exactly the same whether Bernard existed (as a sui generis object) or not, simply because our visual phenomenal experience is caused by light reflected by the objects in front of us, and it would be reflected in exactly the same way whether Bernard were an object or not. The questions surrounding Bernard's existence should then be settled by other than perceptual means; for instance, by invoking philosophical arguments such as Lewis' (1986, p. 212–213) who claims that restricted composition should be rejected because it entails metaphysical vagueness.

There are many other cases of this kind populating metaphysical debates. About the statue and the lump case, for instance, Maclaurin and Dyke (2012, §5) say: "We can talk about whether the statue and the lump of clay are really two objects or actually just one, but the singularity or duality of the statue and the lump is not something that can impinge on human experience." Again, I guess their idea is the same as above: whether there are two objects or only one, our visual phenomenal experience would be exactly the same. The case of our three theories from Chap. 1 can be illustrative here as well: we will not find out in our phenomenal experience whether redness is a trope or a universal or something else, just by looking at a red Ferrari.

The case Merricks (2001, p. 8) is interested in concerns eliminativism about macroscopic objects. Eliminativists claim that there are no apples and no tables, only fundamental components (particles, properties, or whatever one takes to be the

fundamental constituents of reality) *arranged applewise* or *tablewise*. But, the objector could say, we *see* the table and the apple, they're right there, just in front of us. The unfriendly objector could then point to eliminativism as to a doctrine that is almost absurd and entirely in conflict with any good common sense. The less unfriendly objector could instead stress that there is a strong tension between the eliminativist's claim about the world and our experience of the world; the eliminativist, then, at the very least, owes us a good explanation of why there is such a tension, and why we should reject what our eyes see rather than what the eliminativist wants to say. (For a detailed statement of these objections, and replies, see also Sider (2013, especially §5).)

One way to go for the eliminativist could be to show that eliminativism is such a theoretically virtuous doctrine that we should, indeed, reject our intuitions and common sense beliefs about the world based on our experience of apples and tables. Such argumentative strategies are, of course, often used in metaphysics. But the more adequate strategy is to realize that, in fact, there is no tension at all. As before, what we see and what the world is like are just two entirely orthogonal issues, simply because our visual phenomenal experience would be the same whether there were apples and tables in front of us or whether there were 'only' fundamental components arranged applewise and tablewise. Apples and fundamental components arranged applewise both reflect light in the very same way, and, consequently, the way the world 'hits' our perceptual system would be exactly the same in both cases. Thus, our phenomenal experience of the world, which is an experience *as of* apples and *as of* tables, is entirely neutral with respect to the existence of apples, tables, or arrangements of fundamental components. As Merricks puts it, fundamental components arranged applewise can do just about anything apples can do: they can be seen, purchased, eaten, and so on. Thus, there is no tension at all between the eliminativist's picture of the world and our phenomenal experience.

§4. Let us consider in more detail another example: the case of our temporal experience in relation to the A-theory and B-theory of time. In short, B-theory claims that time is very much space-like—a further dimension in which things can be located by bearing relations of 'being earlier than', 'being later than', and 'being simultaneous with' each other—while A-theory claims that time *passes*, that there is a genuine ontological distinction between the present time on the one hand and past and future times on the other hand, and that time is very much *not* like space. Of course, there are many variants of both views, and very different ones[6]; the relevant difference between the two competitors for me in what follows is just that one of them claims that time passes, or flows, while the other does not.

Indeed, the A-theory of time, or at least some versions of it, certainly has as one of its core motivations the idea that it only—as opposed to B-theory—can account for the fact that time *passes* or *flows*, a fact that is taken to be apparent in our experience of time as of passing. In this case, many have thought, and some have

[6]Defenders of one version or another of A-theory include Lowe (1998), Bigelow (1996), Merricks (1999), Markosian (2004), Craig (2000), Zimmerman (1996, 1997, 1998), Prior (1970, 2003b), and Chisholm (1990a, b). Defenders of one version or another of B-theory include Quine (1960), Lewis (1976), Mellor (1981, 1998), Sider (2001), Le Poidevin (1991), and Oaklander (1991).

explicitly expressed,[7] the idea that there is a route from experience to metaphysics, like, for instance, Brad Skow (2009, §4): "I cannot survey all the motivations philosophers have had for the moving spotlight theory. But the motivation that I like best appeals to the nature of our conscious experience". Schematically speaking, the argument goes like this: we experience time as passing (this is given to us in phenomenal experience); B-theory cannot account for temporal passage; A-theory can; thus A-theory is a better theory. A metaphysical conclusion is thus drawn from considerations about our phenomenology. Or, at the very least, the friend of this line of argument says, there is a strong tension between the conception of a B-theoretic 'static' world and our experience of the world which is dynamic—and so the worse for B-theorists.

But even if, perhaps in a more subtle and more controversial way, this case is, in fact, analogous to the two cases we have seen above (about composition and eliminativism) in the sense that issues concerning our temporal experience and issues about the nature of time are, here again, entirely orthogonal. Indeed, as we shall see below, our experience of temporal passage can be accounted for in B-theoretic terms as well as in a dynamic A-world. Our experience of the world is thus neutral with respect to these two theories, and consequently there is no argument from experience to either B-theory or A-theory. Let us see how this works.

The first thing to note is that we actually never experience the flow or passage of time directly, we 'only' experience it *by* having experiences of *succession* and *change*. In agreement with Robin Le Poidevin[8] and L.A. Paul,[9] it seems to me correct to say that if we did not perceive any change at all (including change in our own thoughts), it would not seem to us that time is passing. The way we *measure* time also always involves measuring change: we only measure how much time has elapsed by observing changes happen (for instance, the movement of sand falling through a sandglass).

Now, our temporal experience certainly often is, phenomenologically speaking, an experience as of change, and consequently as of temporal passage. This is often so in perception of *movement*. But such experiences can often be illusory, in the sense that in many cases it appears to us in our phenomenal experience that (i) there is movement when there is not, and that (ii) there is no movement when there is.

Take the case of the hour hand on a mechanical watch: it moves so slowly that we just do not perceive it *as moving*. We can, of course, observe *that* it has moved, if we look at it after some time, but such a "perception" is no experience of

[7]See for instance Williams (1951, pp. 465–466) and Maudlin (2007, pp. 135, 142).

[8]"In fact, it seems odd to say that we see, hear or touch time passing. And indeed, even if all our senses were prevented from functioning for a while, we could still notice the passing of time through the changing pattern of our thought." Le Poidevin (2009).

[9]"[…] if we were in an entirely static environment where there were no contrasts between property instances (this would have to include no contrasts with respect to properties of my thoughts), then it would seem to us as though time were standing still. And, indeed, I think this is a very plausible supposition." (Paul 2010, p. 23).

movement at all. Thus, we have a case where there is movement, but we do not experience it as such, simply because the hour hand's continuous movement is far too slow for us to be able to perceive it. Indeed, what we realize here is that our capacities to notice change and movement have a lower limit and that anything that moves too slowly will not be registered by our perceptual system *as moving*.

Take now the opposite case, where there is no genuine movement but we *do* have a phenomenal experience as of movement. Paul (2010) discusses in detail such cases when she defends, relevantly to our discussion here, an account of how our temporal experience arises from the way our brains interpret cognitive inputs from *series of static events*. To do that, she provides us with some well-known experiments from experimental psychology and cognitive science such as this simple one: "[…] an interesting and empirically well-documented fact about our experience—namely, the illusion we have when, first, one small dot is shown on the left-hand side of a computer screen and then, very quickly, that dot disappears and a small dot is shown on the right-hand side of a computer screen. Then, the right-hand dot disappears, and the left-hand dot appears, again and again, in rapid succession. Even when we are told that what the computer is actually doing is merely blinking different dots on alternating sides of the screen, as long as the succession is rapid enough and spatiotemporally close enough, the effect is that we have the illusion of the dot moving back and forth across the screen. This is what cognitive scientists usually describe as 'apparent motion'" (Paul 2010, pp. 15–16). There are many such examples of apparent motion, not the least of them being the case of cinema where a series of static images is shown quickly enough—but not too quickly—for us to have an experience of movement when we look at it.

What we see here is that a succession of appropriately linked experiences gives rise to an experience of movement, where "appropriately" means that the successive experiences must occur at an appropriate speed, namely, within the lower and upper temporal limits of our perceptual system's capacities—and, crucially, regardless of whether there *is* genuine movement or not. As Paul (2010, p. 16–) argues further, the having of such-and-such experience—as of movement or not—is totally independent of whether we know that we are perceiving a genuinely moving object or whether we know that we find ourselves confronted with an illusion of genuine movement (like in the case of the moving dot on the computer screen or in the case of cinema). The 'apparent motion' illusion persists independently of whether or not we are aware of its illusory nature. The reasons behind all this is that our brain *first* interprets the successive inputs it gets from the world, on a psycho-neurological but not phenomenal level, *before* it produces a conscious phenomenal experience. It 'stores and computes' a number of inputs, and only after a process of interpretation does it give rise to an experience—which will be an experience as of movement or not, independently of whether there was genuine movement at the origin of the inputs or not. This is why we never have an experience of singular static frames in the case of film projection in the cinema: our brain 'creates' for us an experience of movement precognitively before, so to speak, it reaches the phenomenal level of our experience.

Thus, whether we find ourselves in an A-world that contains 'genuine' change and movement, or we live in a B-world that contains a series of static stages which are given to us at an appropriate speed, our phenomenal experience, such as it is produced by our brains and our perceptual systems, will be an experience as of movement (or not, if the A-changes are too slow or too fast, or if the B-series are given to us too slowly or too quickly). As in the case of mereological composition and eliminativism, what we see here is that our phenomenal experience is entirely neutral with respect to what the world is like. No metaphysical conclusions—either in favour of A-theory or B-theory—can thus be drawn from the nature of our experience.

§5. The fourth and last example I want to discuss concerns the way we perceive, and conceive of, the identity of ordinary macroscopic objects through time. The intuitive idea behind *endurantism* is that an apple that is first red and juicy and later brown and rotten is *one and the same* object that is *changing*. Very often this idea is taken to be some sort of philosophically intuitive constraint on the notion of change or the notion of a persisting ordinary object, and endurantism is often claimed—even if not explicitly—to be the 'intuitive' or the 'common sense' view, contrary to perdurantism, as the following citations illustrate.

> No one else [than the perdurantist] would say that only [temporal] parts of Sir Edmund Hilary and Tenzing Norgay climbed only a part of Everest in 1953. The rest of us think those two whole men climbed that one whole mountain, and that all three parties were wholly present throughout every temporal part of that historic event. (Mellor 1998, p. 86)

> […] outside philosophical seminars a four-dimensionalist never says 'a two-hour phase of me last night was a waking phase'; he says, with the rest of us, 'I was awake for two hours last night'. (Simons 2000b, p. 62)

> The [four-dimensionalist's] metaphysics yields that if I have had exactly one bit of chalk in my hand for the last hour, then there is something in my hand which is white, roughly cylindrical in shape, and dusty, something which also has a weight, something which is chalk, which was not in my hand three minutes ago, and indeed, such that no part of it was in my hand three minutes ago. As I hold the bit of chalk in my hand, new stuff, new chalk keeps constantly coming into existence *ex nihilo*. That strikes me as obviously false. (Thomson 1983, p. 213)

In some sense, which is very often not precisely specified, endurantism is supposed to be the obvious, common sense, acceptable, intuitive view, while perdurantism is supposed to violate some strong intuitions we have about the world—or at least this is how I understand the statements that populate the discussion between perdurantists and endurantists. Some perdurantists themselves agree on this "intuitive datum", while they just think that this is a case where we should abandon our intuitions, and revise our common sense beliefs, because they take perdurantism to be a theoretically strong and virtuous view.

Now, what are these pro-endurantist intuitions about? Where do they come from? Are they revealing some philosophically deep conceptual truth? As we shall now see, they are not—rather they are mere consequences of the (contingent) way our perceptual system works, and what's intuitive in the common sense endurantist idea is rather something that arises from the character of our ordinary phenomenal

experience of moving and changing macroscopic objects. In an article stemming from results obtained in experimental psychology experiments, Scholl (2007) discusses what could be an explicit argument in favour of endurantism based on the nature of our phenomenal experience of macroscopic objects:

> "Endurance theories of persistence, for example, may simply strike many readers as much more natural and compelling than perdurance […] theories because endurance theories are more closely matched to the actual way in which we experience the world." […] "Philosophers may think, for example, that they are preferring an endurance theory because 'a number of our practices and forms of self-understanding depend on the idea that there are enduring things' (Haslanger 2003, p. 351)—but I suggest that what many of these 'practices' and 'forms of self-understanding' may come down to in practice are the results of the mental machinery we have that is devoted to computing enduring representations of objects over time." Scholl (2007, p. 583)

The general idea supported by Scholl's findings in the field of experimental psychology is, in short, the following. We want to be endurantists because we have a strong intuition in favour of endurantism, and we have that intuition because of the way we experience the world. But the way we experience the world is (i) contingent (the way our brain and our perceptual system are built—"hard-wired"—is contingent), and (ii) fully compatible with the world being a perdurantist world. As in the other cases we have already seen above (composition, eliminativism, A-theory, B-theory), here again our experience is, in fact, neutral with respect to the way the world is, and so are, then, our intuitions that stem from this experience. (Obviously, this is not to say that endurantism is false—it's just that it cannot be argued for and it should not be motivated by using this experiential and intuitive base.) To see this in more detail, let us consider one of the experiments discussed by Scholl (2007, p. 567) where the way we experience spatio-temporal continuity of moving objects is studied.

In a 'multiple object tracking' (MOT) experiment, observers are first given a series of objects where some of them are marked as being "targets" (by shortly once blinking) (Fig. 6.1).

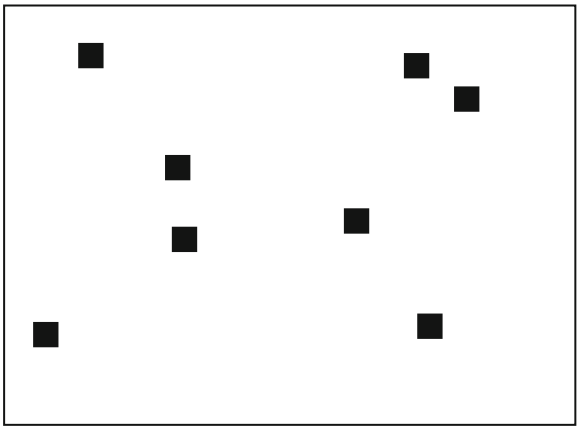


Fig. 6.1 Basic MOT experiment (Animations of this and the following experiment can be downloaded at http://www.yale.edu/perception/Brian/demos/MOT-Occlusion.html)

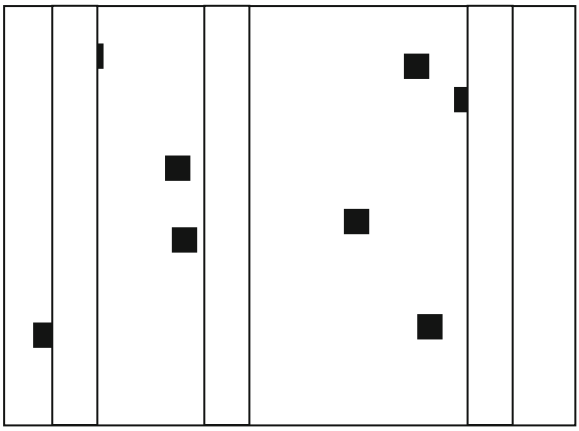


Fig. 6.2 MOT experiment with occluders

The objects then start moving in an unpredictable way in all directions, and when they stop, observers must tell which are the targets and where they are. In a second stage of the experiment, occluders are added to it (Fig. 6.2).

In this scenario, which breaks the spatio-temporal continuity of the trajectory of the moving objects, the targets often disappear behind occluders. As Scholl (2007, p. 567) puts it "objects may not pop into and out of existence, but they do frequently pop into and out of *sight*".

What the results of such (and other variants of) MOT experiments show is that observers are very successful at tracking the targets even in cases where disruptions such as occluders are in place, and that we human beings are just made in such a way that we experience the world as a world of objects that remain the same following a principle of *spatio-temporal continuity*. As Scholl (2007, p. 567) puts it "[the principle of spatio-temporal continuity has] become wired into our minds, and helps to control our experience of objects as persisting in the world". This is not only so when various kinds of occluders are introduced, but even when observers *know* that the targets they are supposed to follow and identify as being one and the same object are in fact *not* one and the same object. Typically, observers are aware of the fact we already encountered in §4 above that the "objects" on the computer screen are actually just a series of different zones successively lit on the screen (to have a more complex and life-like example in mind, just think of tracking an "object" in a movie, where that object often disappears behind occluders, and where you very well know that, in fact, you are viewing a series of static images projected at 24 images per second). In short, it appears that our perceptual system is made in such a way that we track one moving object and identify it as being one and the same object, even if it frequently disappears from our sight, and even if we know that in fact it is *not* one object.

Furthermore, the principle of spatio-temporal continuity appears to be the most fundamental one when it comes to how our experience of persistence works. Indeed, spatio-temporal continuity trumps other criteria, as various other experiments showed. Scholl discusses an experiment where one object disappears behind an occluder and a very different object emerges from the other side of the occluder and continues moving. In such cases, if the second object emerges at the same time as the first object would have emerged had it continued its movement in a continuous way, observers identify it as being one and the same object that underwent qualitative change—even when the change is quite significant (shape, colour, size,…). Our perceptual system tells us that there is *one* and the same object being temporarily hidden by an occluder and undergoing change, rather than two objects one of which stays behind the occluder and the other of which emerges from behind it. Things are different when a temporal delay is introduced between the time when the first object disappears behind the occluder and the time when the second object emerges—in *this* case, where continuity of motion is disrupted, observers identify the objects as being *two* different objects. Thus we see that our perceptual system is made so that it provides us with percepts in such a way that spatio-temporal continuity *trumps* property change—we accept more easily that an object significantly changes its properties, but not so easily that it violates spatio-temporal continuity. To sum up, Scholl (2007, p. 573) says "the results [of these experiments] revealed a *temporal same-object advantage*", and furthermore, what we see in these cases is that we cannot help but having an experience as of one and the same persisting object—it is "forced upon us" by the way our brain and perceptual system deals with stimuli from the external world ("[…] the principle of spatiotemporal continuity […] seems to be wired into our minds in a deep way, controlling how we experience the world" (Scholl 2007, p. 569)).

§6. The results of considerations and experiments such as those we have seen in §4 and §5 above lead, then, to an understanding of our experience of movement, persistence, and diachronic sameness of ordinary material objects as being such that we cannot help but have it, given the way our perceptual system is built, even in cases where there is no motion, no persistence, and no sameness. We have experiences *as of* movement, change, persistence, and diachronic sameness, even where there is none. Thus, as before, we see here that the phenomenal character of our experience is one thing, and the metaphysical nature of the world is another. Whether the world is a perdurantist B-theoretic one or an endurantist A-world, our phenomenal experience would still be the same. Relevantly, we learn from §4 and §5 that we *would* have an experience *as of* persistence of one and the same object identical through time and of time *as* passing, even in a static perdurantist B-world where there is 'only' a series of different numerically distinct objects (temporal parts, or stages) located in a static B-series. And of course, we would have the same experience also in a dynamic endurantist A-world (or in a perdurantist A-world, or

an endurantist B-world, etc.). In short, both a series of successive temporal parts (or stages) in a B-world and an enduring object in an A-world would produce in us, given our perceptual system, the same phenomenal experience.

Now, I would like to suggest that things *also* go the other way around—though, for obvious reasons, *this* claim cannot be backed up here by any empirical research. Indeed, until now I argued in favour of the claim that *our* experience would be the *same* in *different conditions*. But it seems to me also very plausible that *in the same conditions* (in one and the same world) different types of conscious beings, with different perceptual systems, would have *different experiences* of the same metaphysical reality. I will not venture here into a complex discussion about how some animals on Earth experience persistence, movement, temporal passage, or whether they have anything like 'intuitions' in favour of unrestricted mereological composition or eliminativism. But it does seem entirely plausible to me to conceive of alien conscious beings populating some other part of our world who have very different bodies (say, they are made of gas) and very different perceptual systems and brains, and who, as a consequence, have experiences of the same universe that are entirely different from our own.[10]

§7. Gas-based aliens left aside, here are the two claims that I wish to defend: (i) the phenomenal character of our experience is neutral with respect to the nature of metaphysical reality, and (ii) at least some of our intuitions (probably, many) concerning the relevant notions and claims are based on our phenomenal experience and, consequently, are not good guides for the metaphysician to follow.

Let me try to say more about (ii). Indeed, it seems that what is often taken to be an intuition (say, in favour of endurantism, or in favour of A-theory) is simply a result of how we experience the world. We have seen above several examples where our experience seems to back up a metaphysical theory, but we have seen that this is no more than a seeming and a possibly misleading piece of alleged evidence if favour of one or another metaphysical claim. Consequently, one has to be very careful about what one calls an "intuition" and about where such an intuition comes from. If—as I think is indeed the case in the various examples I discussed above—such so-called intuitions are, in the end, just manifestations of how we (contingently) experience the world, then one must take care not to draw any metaphysical conclusions, or to motivate a metaphysical view, based on such 'intuitive data'.

The difficulty is, then, to identify which of the 'intuitions' we have are deceitfully such, and which (if any) are immune to this worry—and this is where work in the field of cognitive science and experimental psychology, as we have seen in the

[10]As illustrations, think of the Dikironium cloud creature, the Beta XII-A entity, or perhaps, Q.

case of Scholl (2007), can be a valuable contribution to the philosophical debate, since it allows us to either discard some of our 'intuitions' as being misleading, or at least to proceed with extreme caution when appealing to them when doing metaphysics.[11]

For the metaphysician, there is a threat there. There is the risk that, at the end of such an evaluative procedure, where our 'intuitions' are scrutinized and their potentially misleading character evaluated, we end up realizing that many, or even most, of what we took to be deep intuitions turn out to be mere constructions from the way we contingently happen to experience the world. This becomes then immediately problematic for metaphysics as a philosophical discipline, since in many domains of metaphysics intuitions play a central role in both theory construction and theory evaluation. Often, in metaphysics, we just don't have at our disposal any other data than our so-called intuitions. What to do?

I try to offer an answer in the next chapter.

[11]In a draft article entitled "Folk Mereology is Teleological", David Rose and Jonathan Schaffer provide another interesting and well-documented case—mereological composition—which bears on the question of the role of (armchair) intuitions in metaphysics: "Our own view is twofold. First, we hold that the folk theory of composition is teleological, in that the folk tend to think that a plurality of objects has a fusion if and only if that plurality collectively serves a purpose. So, for instance, we predict that people will tend to say that composition has occurred with the knife, fork, and plate (since they collectively serve as a table setting) but not with the two people shaking hands (unless they are accorded a function). This view seems not even to be considered in the contemporary discussion, though it coheres with a wide swath of current psychological work on object concepts. Secondly, we regard such a folk theory as tied into a benighted teleological view of nature, and thus deserving dismissal. As such we think that *understanding the folk theory of composition should actually lead us to liberate the discussion of when composition occurs from any demanded conformity with folk intuitions*. On this matter, the folk deserve to be ignored." (my italics).

Chapter 7
The Beauty of Metaphysics

Abstract In this chapter, I put forward the claim that metaphysical theories possess aesthetic properties, grounded in non-aesthetic properties, and that these play a crucial role in theory evaluation and theory choice. The general claim that aesthetic properties supervene on non-aesthetic properties is a largely debated one. In this chapter, I address this issue from an angle which has not really been explored so far: I shall neither concentrate on cases of artefacts nor of natural objects, like the beauty of a painting or the beauty of a sunrise. Rather, my main centre of attention is the somewhat more special, theoretical case of the beauty of philosophical theories (with a focus on metaphysical theories). There are some interesting issues concerning claims that attribute aesthetic properties to theories, in part because, even if such claims are commonplace in philosophy and in science, little has been said about the nature of the relevant supervenience base—that is, about what it is exactly that the beauty of a theory is supposed to supervene on. The relation of supervenience itself is questioned. We shall see that aesthetic properties of theories play a crucial role in theory choice and evaluation. Finally, I offer reasons to think better of a kind of anti-realism that the claims I have been defending in this book appear to lead to.

§1. Let me sum up where we find ourselves. We have seen, in Part I, how important primitives (primitive problem-solvers) are, and how they play a crucial role in the construction and evaluation of metaphysical theories. This made me ask how to choose between competing metaphysical views. In some cases, the answer is that we don't have to make a choice since the two allegedly competing views turn out to be equivalent. But in many other cases this is not so, as for instance in the case of trope theory, the theory of universals, and resemblance nominalism, which I used as my main example in Chap. 1. The point of Chaps. 5 and 6 was, then, to realize that even in such clear cases of non-equivalence (on purpose, I used here an uncontroversial example of non-equivalence, instead of a controversial one, like the debate between eternalism and presentism discussed in Chap. 4), it is far from being clear on which grounds we could select one theory as being better than another. I critically discussed in detail two main criteria for theory evaluation (explanatory power (in Chaps. 4 and 5), and intuitions (in Chap. 6)), as well as criteria like simplicity, parsimony, internal consistency, and others

J. Benovsky, *Meta-metaphysics*, Synthese Library 374,
DOI 10.1007/978-3-319-25334-3_7

(in Chap. 5), with the result that such tools are not able to play a decisive role in theory evaluation and theory choice. So, the naïve (but I believe genuine) question I raised in Chap. 5 remains: *why is it that we metaphysicians don't all end up defending the same view*? Rescher (1978, pp. 218–219) nicely puts it: "why are [philosophers] chronically incapable of reaching a meeting of minds?" We all know the same arguments, we all have in our minds the same lists of pros and cons when it comes to the theories we work with, we are all rational and we share the standards for argumentative work—but still, some of us turn out to be trope theorists, or friends of universals, or resemblance nominalists, to stick to my example from the preceding chapters. Perhaps we sometimes make the mistake of appealing to alleged (misleading) intuitions, in the way I critically discussed in Chap. 6. But even when we don't, it seems that something else is at stake, and that something else plays an important role in the way we choose one theory as being better than its competitors. In this chapter, I will be interested in what this something is.

§2. In a nutshell, my claim will be that metaphysical theories possess aesthetic properties, grounded in non-aesthetic properties, and that these play a crucial role in theory evaluation and theory choice. The practice of attributing aesthetic properties to scientific and philosophical theories is commonplace. Perhaps one of the most famous examples of such an aesthetic judgement about a theory is Quine's in 'On what there is' (1948): "Wyman's overpopulated universe is in many ways unlovely. It offends the aesthetic sense of us who have a taste for desert landscapes […]". Many other philosophers and scientists, before and after Quine, have attributed aesthetic properties to particular theories they are defending or rejecting. One often hears that a view is "elegant", "attractive", "beautiful", or even "sexy". The physicist Brian Greene decided to call the book where he explains and defends the theory of superstrings for a general readership "The elegant universe". And Dirac commented on general relativity theory thus: "The foundations of the theory are, I believe, stronger than what one could get simply from the support of experimental evidence. The real foundations come from the great beauty of the theory. […] It is the essential beauty of the theory which I feel is the real reason for believing in it" (Dirac 1980, p. 10). When defending four-dimensionalism, Ted Sider could not be more explicit: "It is easy to feel […] an intellectual joy in contemplating a theory so elegant and beautiful as four-dimensionalism, and it is tempting to accept the theory simply on this base, utilizing arguments to rationalize more than justify" (Sider 2001a, p. 74). The list could go on and on.[1]

The general claim that aesthetic properties supervene on non-aesthetic properties is a largely debated one.[2] In this chapter, I will address this issue from an angle which has not really been explored so far: I shall neither concentrate on cases of artefacts nor of natural objects, like the beauty of a painting or the beauty of a sunrise. Rather, my main centre of attention will be the somewhat more special,

[1]Derkse (1992) and McAllister (1999) feature a number of relevant quotes and references.

[2]See, for instance, Sibley (1959), Levinson (1984), Pettit (1987), MacKinnon (2001), Zangwill (1994, 2001, 2003).

theoretical case of the beauty of philosophical *theories* (with a focus on metaphysical theories). As we will see, there are some interesting issues concerning claims that attribute aesthetic properties to theories, in part because, even if such claims are commonplace in philosophy and in science, little has been said about the nature of the relevant *supervenience base*—that is, about what it is exactly that the beauty of a theory is supposed to supervene *on*.

First, as we shall see, supervenience is actually not the right kind of relation that holds between aesthetic properties and non-aesthetic properties of theories. Rather, *grounding* is a better candidate (this will connect this discussion to the claims I raised about primitives in Chap. 4).

Second, in §8–9, I will discuss the role that the context of origin of a theory and the taste of the evaluator play in the attribution of aesthetic properties to metaphysical theories. These two elements will provide an answer to my naïve question about why we don't all end up defending the same views.

Third, in §10, we shall see that aesthetic properties of theories play a crucial role in theory choice and evaluation. Indeed, Sider's and Dirac's quotes above already anticipate explicitly a point I will discuss in detail: it seems that the aesthetic properties of a theory can be appealed to when it comes to preferring one theory over another.

Finally, in §11, I will offer reasons to think better of the kind of anti-realism that the claims I have been defending in this book appear to lead to.

Aesthetic Supervenience, Aesthetic Grounding

§3. The claim that the having of aesthetic properties supervenes on the having of non-aesthetic properties has been widely discussed and, in various ways, defended (see, *inter alia*, Sibley 1959; Levinson 1984; Pettit 1987; MacKinnon 2001; Zangwill 1994, 2001, 2003). But, as we shall now see, even if it is sometimes true that a *supervenience* relation holds between aesthetic properties and the 'subvenient' non-aesthetic ones, it is not the *interesting* relation in the neighbourhood. A richer, asymmetric and irreflexive relation is required, and I shall defend the claim that the more-and-more-popular relation of *grounding* does a much better job than supervenience.

The core idea behind the claim of aesthetic supervenience is philosophically appealing as well as intuitively adequate: aesthetic properties, such as the property of being beautiful, elegant, or ugly are exemplified by the objects that possess them in virtue of having other non-aesthetic properties such as colour distribution or shape. The beauty of a painting is thus said to depend on how paint is arranged on a canvas, while the beauty of a sunrise is said to depend on an arrangement of clouds, rays of light, colours, and so on. I have nothing to say against this core and intuitive claim, and I think that it is correct. There is, indeed, a relationship between the aesthetic and non-aesthetic properties of an object, where the latter are somehow the base for the having of the former. The *in*correct claim that I do have something to say against is that this relation is *supervenience*.

Supervenience is a relation well known from many philosophical debates, including philosophy of mind, metaphysics, and ethics. It is standardly defined as follows (Bennett and McLaughlin 2005): "A set of properties A supervenes upon another set B just in case no two things can differ with respect to A-properties without also differing with respect to their B-properties. In slogan form, 'there cannot be an A-difference without a B-difference'."

Given the serviceability of supervenience in many philosophical debates, the idea of appealing to the same kind of relation in aesthetics to explain how aesthetic properties supervene on non-aesthetic ones does indeed look appealing, and has been largely influential and widely defended. Zangwill (2001, p. 43), for instance, defends this claim ("[…] a supervenience relation holds between aesthetic and nonaesthetic properties"), as well as, famously, Levinson (1984, p. 93) claims that "[…] aesthetic attributes of an object are supervenient on its non-aesthetic ones. […] I will take the thesis of aesthetic supervenience to be roughly this: […] Two objects (e.g. artworks) that differ aesthetically necessarily differ non-aesthetically."

To make these claims plausible and informative, it is necessary to supplement them with more data about the nature of the supervenience base—the question is, what exactly do the aesthetic properties of an object supervene on? Indeed, Levinson, Zangwill, and many others all agree that in addition to colours, shapes, and the like, some *relational* properties also have to be included in the 'supervenience base'. These include, typically, the history and context of *production* of an artwork (more on this in §8 below) as well as the context of *evaluation* (more in §9). If one agrees to thus enrich the supervenience base with the context and history of production, as well as the evaluator's disposition to certain responses, we now have a claim of aesthetic supervenience accurately and completely articulated. Now, let us see where trouble lies.

§4. In short, the trouble is that supervenience is not a relation that is suitable to play the theoretical role we want it to play here. Some think that this is not important, and that it is enough to have a loose and imprecise intuitive understanding of the relationship between aesthetic and non-aesthetic properties in order to defend a claim of 'aesthetic supervenience'. This is the case, for instance, for Zangwill (2003, §2) where he says:

> A fundamental principle is that aesthetic properties are *determined by* or are *dependent on* nonaesthetic properties. Things come to have aesthetic properties *because of* or *in virtue of* their nonaesthetic properties. For example, a performance of a piece of music is delicate *because of* a certain arrangement of sounds, and an abstract painting is brash or beautiful *because of* a certain spatial arrangement of colors. In the philosophical jargon, aesthetic properties *supervene* on nonaesthetic properties. This means that if something has an aesthetic property then it has some nonaesthetic property that is sufficient for the aesthetic property. (The relation of dependence or supervenience is a general one. I shall not probe the exact nature of the relation, although it can be formulated in different ways. The notion is important outside of aesthetics, in areas like moral philosophy and the philosophy of mind.) (original italics)

Zangwill seems here to be well aware of the fact that there are many different types of relations in the neighbourhood of supervenience, and he just appears to think that it is not important for the aesthetician to decide which one best suits the

relationship between aesthetic and non-aesthetic properties (at least not in the context of his Zangwill (2003)). He thus speaks about this relationship as being one of *supervenience* (see also Zangwill (2001, p. 43): "[…] a supervenience relation holds between aesthetic and nonaesthetic properties"), or one of "metaphysical *determination*" (see also Zangwill 2003, §3), *dependence*, or *explanation*.

My intent here is not to criticize Zangwill's own view (which, again, I find largely philosophically appealing and intuitively adequate), but rather to bring forward my concern that not enough attention has been paid in the literature to the exact nature of the relation, often called 'supervenience', between aesthetic and non-aesthetic properties of an object. Levinson (1984, p. 94), for instance, also says interchangeably that aesthetic properties *supervene* on non-aesthetic ones, or that aesthetic differences are *grounded* in non-aesthetic differences, or that non-aesthetic properties are *responsible for* aesthetic ones. But all of these relations are different, they exhibit different formal features, and they are suitable to play different, incompatible, theoretical roles. Thus, in order to simply *understand* what the claim of 'aesthetic supervenience' amounts to, we have to understand not only the nature of the supervenience base (which *has* been discussed at length and in detail in the relevant literature) but also the nature of the relation between such a base and the properties that are said to arise from it. Let us start by seeing why *supervenience* can*not* do the job.

Remember that supervenience is typically defined as follows: "A set of properties A supervenes upon another set B just in case no two things can differ with respect to A-properties without also differing with respect to their B-properties. In slogan form, 'there cannot be an A-difference without a B-difference'" (Bennett and McLaughlin 2005).

A first problem that arises when one wants to apply this notion of supervenience to the relationship between, say, the beauty of Dalí's *Temptation Of Saint Anthony* and its non-aesthetic (intrinsic and relational—see above) properties is that, upon the very same properties (the very same 'supervenience base') supervene also, with the same sort of necessity, *many other non-aesthetic properties* of Dalí's painting, and consequently the claim of aesthetic supervenience fails to single out what's special about the way the *aesthetic* properties of the painting arise from their supervenience base. This is simply because, given the nature of the relation of supervenience, necessary properties of any object supervene trivially on any properties it has (see, for instance, Correia 2008; Bennett and McLaughlin 2005). Typical stock examples include properties such as being self-identical, which any object has necessarily, and since it has them necessarily it cannot differ with respect to them, and so it cannot differ with respect to them without differing with respect to any other property the object has. It is impossible for Dalí's *Temptation Of Saint Anthony* not to be self-identical, and so it is impossible for this painting not to be self-identical without, say, featuring four elephants. Thus the property of being self-identical supervenes on the property of featuring four elephants, or any other of the painting's properties, and relevantly to our present concerns, it supervenes on the very same set of properties (paint distribution, canvas size, history and context of production,…) on which supervenes *the beauty* of *The Temptation Of Saint Anthony*. Since the very same relation of supervenience thus holds between aesthetic properties and their supervenience base, *and* between many

other (necessary) properties of the same object and the very same supervenience base, the claim of aesthetic *supervenience* becomes trivial and uninformative when it comes to understanding the nature of the relationship between *aesthetic* and non-aesthetic properties. To make it perhaps clearer: the trouble here is that we want to say that the painting has its aesthetic properties in virtue of having its non-aesthetic properties, but saying that this is so because its aesthetic properties supervene on its non-aesthetic ones is of no help at all, since supervenience holds in a very trivial way even between totally unrelated properties, such as the property of being self-identical and the property of featuring four-elephants.

(Note that this problem with necessary properties extends to the having of 'impossible properties' as well, since no object can differ with respect to properties that no object can possibly have, like the property of being a talking donkey *and* not being a talking donkey, and so no object can differ with respect to the (non-)having of such properties without differing with respect to (any) other properties—thus the non-having of the property of being a talking donkey and not being a talking donkey, that Dalí's painting (as any object) trivially doesn't have, supervenes on the very same properties (paint distribution, canvas size, history and context of production,…) on which supervenes the painting's beauty).

§5. A second type of worries arise with respect to some formal features of the relation of supervenience. When saying that aesthetic properties supervene on non-aesthetic properties, what we want to say is that aesthetic properties arise from the non-aesthetic ones (the supervenience base) in the sense that the latter are somehow *prior* to the former—we have here the idea that there is an *order*. But supervenience does not have the right features that could provide us with the desirable order, since it is not a priority relation (see, again, Bennett and McLaughlin 2005). Indeed, nothing in the nature of the relation of supervenience says anything about which (if any) of the two terms of the relation is prior to the other.

Firstly, as Kim (1984) noted, supervenience is a *reflexive* relation: "[…] To see that supervenience is reflexive, note that for any set of properties A, there cannot be an A-difference without an A-difference". But nothing can be prior to itself, and thus the relation between the aesthetic and the non-aesthetic properties of Dalí's *Temptation Of Saint Anthony*, which *is* a relation of priority, is not reflexive—and, consequently, it is not supervenience.

Secondly, a relation of priority, such as the one between aesthetic and non-aesthetic properties of a painting, is an *asymmetrical* relation: if a set of properties is prior to another, then the latter cannot be prior to the former (nothing can be prior to anything that is prior to it). Supervenience, however, is neither symmetric nor asymmetric. It can be symmetric, since any case of reflexive supervenience is trivially a symmetric one, and, more controversially, it can also be said to be asymmetric for instance if one thinks that mental properties supervene on physical properties, and that mental properties are multiply realizable, which means that physical properties do *not* supervene on mental properties.

§6. Far from being a logician's or metaphysician's hair-splitting quibbles, unimportant for the aesthetician, the technical concerns discussed in the last two sections show us that something more than 'mere supervenience' is required to help

us understand the nature of the relation between aesthetic and non-aesthetic properties. Indeed, supervenience just is not the right kind of relation to play the role we need it to play. As McLaughlin (1995) rightly notes, a supervenience claim does not automatically entail an "in virtue of" claim. As we have seen above, supervenience is not a relation of priority, and it just does not say anything genuinely informative about the relationship between aesthetic and non-aesthetic properties—it does *not* say that an object has its aesthetic properties *in virtue of* or *because* its non-aesthetic ones, since it can generally be the case that a set of properties of *x* supervenes on another set of properties of *x* while *x* does *not* have the former in virtue of having the latter. (Trivially, for any object, and for any property F, *x*'s having F supervenes on its having not-F since an object cannot differ with respect to F without also differing with respect to not-F, but clearly an object is not said to have F in virtue of having not-F).

We need more than just supervenience. What we need is a relation that links aesthetic and non-aesthetic properties of an object in a more intimate manner—in a way which *makes sense* of Dalí's *Temptation Of Saint Anthony* having such and such aesthetic properties *because* it has such and such non-aesthetic (intrinsic and relational) properties. Its having these non-aesthetic properties must somehow *explain* that it has the aesthetic ones. Supervenience is merely a form of *covariation* of (sets of) properties,[3] but lacks any element of explanation about why such covariations occur. As Kim (1984) puts it: "[…] we look at the relationship as specified in the definition between a strongly supervenient property and its base property, all that we have is that the base property entails the supervenient property. This alone does not warrant us to say that the supervening property is dependent on, or determined by, the base, or that an object has the supervening property in virtue of having the base property."

Thus, the thesis of aesthetic supervenience, when understood as a thesis of aesthetic *supervenience*, is rather frustrating and uninteresting. Supervenience is not a relation of priority, it is merely covariation, it does not have the right formal features, and it cannot do the job that we need it to do. But *grounding* can.

§7. Aesthetic grounding is what we need, instead of aesthetic supervenience. The relation of grounding, more and more popular in metaphysics, is also one that we can appeal to in order to understand the relationship between aesthetic and non-aesthetic properties.

Grounding cannot be analyzed successfully in terms of supervenience or in other terms (see Schaffer 2009, forthcoming) and it is best taken as being a primitive relation. But the fact that it is a primitive does not prevent us from explaining (instead of defining) and understanding what it is, and putting it to good use. The core idea behind grounding is an Aristotelian idea about a structured nature of the world, as Schaffer (2009) argues. This is, perhaps, why it is of the utmost interest for the metaphysician. Metaphysics studies the nature and structure of being, and as Schaffer and many others claim, such an enterprise is best seen as trying to find out "what

[3]See Correia (2008): "The concept of supervenience involves the notion of 'covariation': that which supervenes (the set of supervenient properties) 'covaries' with that on which it supervenes (the set of subvenient properties), i.e. there can be no 'variation in' the supervenient properties without some 'variation in' the subvenient properties."

grounds what", rather than merely "what there is". The main idea here is that, rather than just looking for a *list* of what exists, we'd better (also) try to see what is more *fundamental* than what—what grounds what—giving rise to a view about the nature of the world as being a hierarchical structure (rather than a list). Under this conception of the metaphysician's job, in Schaffer's terms, one is not as much interested in what exists as in *how* things are, and one tries to see which entities are fundamental (i.e. 'grounds') and which are derivative (i.e. 'groundeds'). It is not that existence questions are to be overlooked, rather the claim here is that answering them does not constitute the ultimate goal of metaphysical enterprise. Typical examples of grounding relationships include not only metaphysical claims such as "the temporal extension of a material object is grounded in its temporal parts" or "a trope is grounded in its bearer" (see Correia 2008), but also claims like "moral features are grounded in natural features of a situation" (see Schaffer forthcoming), and importantly for us, claims like "aesthetic facts are grounded in non-aesthetic facts" which Audi (forthcoming) cites "among the most compelling examples [of grounding]".

One can directly see that the relation of grounding does not have the frustrating features of the relation of supervenience. If a set of properties A is grounded in a set of properties B, and if B is more fundamental than A, then we not only secure the claim that there cannot be an A-difference without a B-difference, but we also have an answer as to *why* an object has A when it has B. That is, grounding *is* a relation of *priority and ordering*, it is irreflexive and asymmetrical, it does not yield the unwelcome result that necessary properties are grounded in anything, and it does not force us to say that being F is grounded in being not-F. This is how we want the relationship between aesthetic and non-aesthetic properties of Dalí's *Temptation Of Saint Anthony* to be. Indeed, while the grounding relation works with the same 'base properties' of the painting as the supervenience relation (including the *broad* non-aesthetic relational properties, like the process and history of production of an artwork as well as the context in which it was created, etc. (see §2)), it does not amount to a mere covariation between the higher-order properties and the 'base' ones. The concern here is not only to use proper terminology—and to stop using "supervenience" as a term for talking about the relation between aesthetic and non-aesthetic properties—but, *by* using proper terminology, make it clear what features such a relation has to have and what theoretical role we want it to play. Understanding this thesis as a thesis about grounding gives us, then, a better understanding of what the relation between aesthetic and non-aesthetic properties is, and what it is *not*.

One might object that, granted, the grounding relation has all the right formal features we need it to have in order to play the theoretical role in linking non-aesthetic and aesthetic properties of an object, but that, being a *primitively postulated* relation, it does not tell us much—remember the concerns I raised about primitive problem-solvers and explanatory power in Chap. 4. To paraphrase Locke's worries about substrata, grounding is a "we-know-not-what". We know, at least to some extent, what grounding *does*, or what it is supposed to do, but we are not told much about its nature. A deeper account of what grounding is, of why it holds between non-aesthetic and aesthetic properties of an object, and of what kind of explanatory power its obtaining gives us, seems to be required.

In one sense, these worries cannot be answered. Indeed, if grounding really is a primitively postulated tool, introduced by the philosopher who realizes that supervenience cannot do the job and that something else must do it, then it is, in some sense, impossible to ask for more: as we know from Part I, one cannot dig deeper than a primitive. Primitives are the pillars that sustain the architecture of our philosophical theories, and it is, in some sense, unfair to ask one's opponent to go beyond what she takes to be a non-analyzable central feature of her view. But then, the worry becomes different and perhaps even more pressing. Indeed, the theoretical job of the grounding view, as proposed above, is achieved through a primitive. (This is actually also the case for the supervenience view.) But then, the questions I raised in Chap. 4. §5–§6 arise here again: where does the explanatory power of such a theory come from? Is all or almost all of the view's explanatory power just primitively postulated? What and how exactly does grounding *explain* here? How can a primitive explain anything?

The answer here is the same as before: granted, both supervenience and grounding (as well as many other primitive problem-solvers) are primitives, and in a sense we frustratingly cannot go deeper than them. But this does not mean that we cannot understand the ways in which they play their theoretical role. Both candidates for being the best relation that obtains between non-aesthetic and aesthetic properties of an object are primitive "we-know-not-what" but they *are* "we-know-what-it-does" (remember the functional view, discussed in Chap. 4, §2–§4), and the purpose of the preceding sections was to establish that grounding is a more serviceable hypothesis than supervenience. It has the right formal features, it is a relation of explanation, priority, fundamentality—it is an 'in virtue of' relation known to be a useful hypothesis in many other areas of philosophy, and, consequently, it does a better job than supervenience in the case of aesthetic properties. It does not tell us much about itself, but it does tell us a bit about its relata: as Bricker (2006), De Rosset (2010), and Schaffer (2009, manuscript) claim, if a is grounded in b, a is nothing over and above b. a, in other words, is an "ontological free lunch" in Armstrong's (1997) sense; the "ontological price", to use Schaffer's (manuscript) term, you pay for a and b is just whatever you would pay for b alone. Only in this sense can one talk about identity between a and b. This does not mean that aesthetic properties are not real, no more and no less than lightning or than two objects' sharing the same property or than an object's persisting through time. Aesthetic properties, on this view, are as real as the non-aesthetic ones they are grounded in, it's just that the latter are the more fundamental ones.

The Context of Origin

§8. As already mentioned above, and just as it is often argued in the case of artworks, the base in which aesthetic properties of theories are grounded should include the context within which a given theory was formed. Exactly as the aesthetic value of a painting or a novel is influenced by its creative origins (the historic, social, political, etc. contexts determining the artwork's originality or even meaning

(for instance, in the case of a novel such as Orwell's "1984")), the respective historical period and the context of the creation of a philosophical theory will matter for the attribution of aesthetic properties.

More precisely, in addition to colours and shapes, etc.—but note that different types of objects, like symphonies, would possess other first-order non-aesthetic properties—some *relational* properties also have to be included in the base on which aesthetic properties of artworks are grounded. These include, typically, the history and context of production of an artwork (see, for instance, Levinson 1984, p. 93–94). In Walton's (1970) terms, aesthetic properties of an object depend not only on its *narrow* non-aesthetic properties, but also, importantly, on *broad* non-aesthetic relational properties, like the process and history of production of an artwork as well as the context in which it was created. Two indistinguishable paintings, indiscernible in the sense that they are exact duplicates and exactly the same arrangements of paint on a canvas of the same size, shape, texture, and so on, would (or, at least, could) still possess *different* aesthetic properties depending, for instance, on the period when they were created.

This 'broadening' of the grounding base solves a problem raised by Scruton (1974, p. 36), who criticizes the aesthetic supervenience thesis when he says that "different emergent 'properties' can depend on precisely the same set of 'first order' properties". What he has in mind here is that one and the same artwork can be context-dependently characterised as sad or as joyful, without contradiction. (For a discussion of this phenomenon, see for instance Pettit 1987; Zangwill 1994; MacKinnon 2001.) We can now respond to such an objection simply by pointing out that, once we include the context of production (*and* the context of evaluation—see more on 'taste' in §9 below) in the grounding base, it is *not* anymore the case that 'different emergent properties could arise from the same base'.

When it comes to theories, the kind of relevant context I have in mind is, in particular, the state of philosophical and scientific knowledge at the time of the formation of the theory to be evaluated. Take the case of Thales' materialist conception of the world, based on the idea of water as the central element out of which all other existing material entities are somehow construed. Clearly, such a view, evaluated in the light of today's scientific and philosophical knowledge, is false and not very satisfactory with respect to several of the evaluative criteria we have seen in Chaps. 4–6 (explanatory power, compatibility with other successful theories, etc.). Does this mean that Thales' view cannot be said to be beautiful? No, for the reason mentioned above: the context of origin of this metaphysical theory is to be taken into account when evaluating the theory's beauty, exactly as in the case of works of art. This means that, when we say that aesthetic properties of theories are grounded in their non-aesthetic properties, the grounding base has to be widened to include their context of origin as well as the other non-aesthetic features—and, from the point of view of scientific and philosophical knowledge in the 6th century B.C., Thales' theory represents quite an achievement, in terms of systematization and philosophical reflection.

But this approach seems to generate a result that, while welcome in the case of artworks, such as paintings, is distinctly undesirable in the case of metaphysical theories: Thales' view (and, of course, many an ancient, medieval, and modern

view) could very well emerge from the evaluative procedure as being judged just as beautiful as the best metaphysical theories we have today. This state of affairs is acceptable in the case of, say, paintings since there is no good reason for claiming that today's paintings are in any principled way superior to older ones, but it is an unacceptable result in the case of philosophical theories, because it does not do justice to the progress of philosophical knowledge. In short, we want to say that, even if it is not always the case, generally speaking, our theories become better—more beautiful—over time (recall Sider's and Dirac's quotes; indeed, all this becomes crucially important if one takes the *beauty* of a theory to drive one's choice in deciding which particular theory is supposed to be the *best*—more on this below). But it seems that, if aesthetic properties are grounded not only in their intrinsic non-aesthetic features but also in a wider base that includes the context of origin, it could perhaps even be possible to judge Thales' view as better (because more beautiful) than some of the most elaborate theories we have today.

However, the impasse here is only apparent. For, unlike paintings or other art forms, metaphysics exhibits one important feature which it shares with all the other sciences and philosophical disciplines: its knowledge accumulates over time. Another way of bringing out this point is to say that the contemporary context of origin of metaphysical theories does, in a certain sense, include all past contexts, since it includes all the successful discoveries of the past. This is why the contemporary context is to be privileged over any other past contexts, and, consequently, contemporary metaphysical theories can be said to be better than past ones (if they are beautiful enough) and claims about the progress of knowledge in philosophy can be secured.

Nevertheless, the general idea I wish to propose here still *is* analogous to the case of artworks like paintings. Its core claim is simply this: since the context of origin is part of the base in which aesthetic properties of theories are grounded, the context of contemporary theories is *richer* than the context of 'older' theories. Suppose I see a painting in the museum which I intuitively like and find beautiful without however knowing anything about its context of creation. Suppose further that a museum guide comes along and provides me with interesting background information about the relevant context, for instance, that the painting was created in the Czech Republic in the seventies and that it has a particular political significance as a metaphorically veiled rejection of the communist regime at the time. After I have been given this information, I might find the painting even *more* beautiful than before. Suppose the guide goes on to tell me about the painter's life and reveals to me even more about the context of the painting's creation, for instance, that the painting also offers a metaphorical reference to the day when the painter lost his child—I might again find the painting now *even more* beautiful. In short, what I want to express here is the general thought that, the richer the context, the (potentially) more beautiful the painting will be to the beholder. What is more, it seems to me that this applies even more clearly in the case of philosophical theories (remember here the "widen the net" criterion I discussed in Chap. 5, §7).

The Taste of the Evaluator

§9. The taste of the evaluator and her capacity to recognize and assess aesthetic features of works of art has often been argued to be of the utmost importance in the attribution of aesthetic properties; indeed, it seems that it should be included in the base in which the aesthetic properties of the artwork are grounded. Here, 'taste' does not simply stand for 'liking' but a more elaborate capacity of the evaluator (one that can be trained), as for instance Sibley (1959, p. 423) makes clear: "When I speak of taste […], I shall not be dealing with questions which center upon expressions like 'a matter of taste' (meaning, roughly, a matter of personal preference or liking). It is with an ability to notice or discern things that I am concerned."

The interesting difference between the case of artworks (paintings, novels, etc.), on the one hand, and the case of philosophical or scientific theories, on the other, is that the claim is much *less* controversial in the latter case than in the former. Consider Hume's assertion that not everybody's taste provides for a good enough judgement, that is, the claim that not everyone is a good art critic (see Hume's 1985 'On the standard of taste'). When it comes to evaluating philosophical and scientific theories, such a statement amounts to something quite trivial: only trained and informed philosophers and scientists can claim to be good judges of the beauty of theories. Furthermore, following Hume (1985, p. 240–241), these qualified judges must obey additional constraints, such as, being practised in the attribution of aesthetic properties to theories, having a "good sense", and being intellectually honest (for instance, in avoiding both jealousy and sympathy towards the author when evaluating her theory). While in the case of Hume's view, such a claim is of course controversial, since he was interested in aesthetic judgements about works of art where *prima facie* anybody feels that she 'has the right' to claim to be a good judge of what is beautiful and what is not (see Sibley (1959) for an interesting discussion), in the more limited case of evaluating philosophical and scientific theories, disqualifying untrained 'common sense' opinions only seems the natural thing to do—indeed, the beauty of theories is grounded at least partly in their non-aesthetic features and only if those features are known and well understood can one start to be a good judge of the overall beauty of the theory at hand.

Much less controversially than in the case of artworks like paintings, then, it seems true that (i) only the aesthetic judgements of trained, qualified, relevantly competent, and appropriately sensitive and receptive philosophers and scientists, who exercise their taste in a proper way, should count, and that (ii) these judgements should indeed be included in the base in which a theory's aesthetic properties are grounded. In short, only a trained and perceptive philosopher or scientist will be able to notice and appreciate a theory's beauty, and her taste and judgement is crucially relevant to any attribution of aesthetic properties.[4] (In the next section, I shall say more about the *role* taste plays here.)

[4]There are two possibilities here. First, one can include taste directly in the grounding base, which makes the aesthetic properties of theories response-dependent (that is, no appreciators, no aesthetic properties); or second, one can include taste only as a condition for the recognition of aesthetic

Theory Evaluation and Theory Choice

§10. We now have a better idea of the nature of the base in which aesthetic properties of theories are grounded. We also already saw (recall the quotes from Dirac and Sider above, to which one could add many other examples from the history of philosophy and science) that many evaluative aesthetic judgements proffered by skilled practitioners of philosophy and science aim not only at the attribution of aesthetic properties to theories for their own sake, but also aim at providing a base for *choosing one theory over another*.

It seems to be a natural attitude amongst scientists and philosophers to be inspired by the beauty of a theory in defending it against its competitors. Beauty (or other aesthetic properties of theories, such as, elegance) thus seems to be an important meta-theoretical criterion when it comes to theory choice. If this were indeed so, aesthetic values would render us a great service because, very importantly, as we have seen in Chaps. 4–6, none of the traditional evaluative non-aesthetic criteria can in fact assist us in selecting one theory over another (remember the discussion about internal consistency, explanatory power, simplicity, parsimony, preservation of and compatibility with intuitions, compatibility and fruitful interaction with other philosophical and/or scientific theories, etc.).

Does this mean that I recommend full rejection of the use of intuitions in metaphysics, or of considerations about a theory's simplicity and parsimony, of its explanatory power, and other meta-theoretical criteria? No, since, as we have seen, they can be useful at least to some extent in some debates, but also because even if one of these criteria alone cannot do the job of telling us which theory is the best, the combination of several (or all) of the various meta-theoretical criteria could perhaps have the cumulative effect of selecting one candidate as being better than the others. But what would it mean to 'cumulate the effect' of these criteria? If the effect of one is such that it does not allow to give clear preference to one candidate, how can the effect of two, three or more such criteria 'become' decisive? Where exactly would the decision-making power come from?

This is where the view (i) that philosophical theories possess aesthetic properties such as "being elegant" or "being beautiful" which are partly grounded in the non-aesthetic meta-theoretical criteria such as those we have seen in Chaps. 4–6, and (ii) that the attribution of these aesthetic properties plays a crucial role in selecting one theory as being better than the others, becomes of great service. In short, the view at hand is that philosophical theories are beautiful and that contemplating their beauty is what drives us to prefer one to another.

The beauty of a sunset is grounded in its non-aesthetic properties such as having this or the other chromatic properties, and/or this or the other pattern of clouds, and

(Footnote 4 continued)

properties. While I have sympathies with the first option, my main point does not depend on it: one can recognize the utmost importance of the role taste plays in the attribution of aesthetic properties to theories even under the second reading.

so on. The beauty of philosophical theories is grounded in their non-aesthetic features such as those put forward by the various meta-theoretical criteria (internal consistency, explanatory power, capacity for being included in a wider network of human knowledge, compatibility with one's intuitions, and of course the various kinds of simplicity and parsimony). Thus, the claim that selection of the best theory is done in virtue of its aesthetic properties does *not* constitute an *additional* meta-theoretical criterion on a par with the others, rather it makes this new criterion to be the cumulative effect of some or all of the non-aesthetic meta-theoretical features of a theory. It does *not* amount to the *abandonment* of all these meta-theoretical criteria, on the contrary, it *makes sense* of the way they function in the process of evaluation and selection of a candidate as being the best.

One way to see in what sense the claim that aesthetic properties of theories *do* allow us to make a choice, combined with the claim that the 'first-order' criteria enter into the grounding base for the 'second-order' aesthetic properties of competing theories is to ask—again—my naïve question: if, first, we are all able to attribute properties like 'exhibiting such and such a kind of simplicity' and similar to our theories, and, second, we all have the same 'first-order' data (that is, we all know all the pro and con arguments there are for each theory, we know how they stand with respect to all of the meta-theoretical criteria mentioned above, etc.), and, third, we all work in a rational way—*why is it that we don't all end up defending the same view*?

There can be many reasons for philosophical disagreement, some philosophical, and some perhaps psychological or sociological (to have an article published and help one's academic career, it's easier to disagree with one's opponents than to 'simply' agree with the existing literature, …). But the generally relevant answer we can give here, keeping in mind all of the considerations we have taken into account in the preceding sections, is simple enough: some of us are more receptive to the beauty of desert landscapes, while others prefer the varied beauty of 'urban landscapes', such as the crowded centres of skyscraper-filled big cities. Some feel attracted by Bauhaus simplicity, while others by Baroque complexity. Some of us feel aesthetically stimulated by simplicity of structure, yet others are more inclined towards simplicity and parsimony with respect to basic axioms of a theory. Some of us are struck by the elegance of a view that shows great explanatory power, while others feel more attracted towards a theory that preserves one's pre-theoretical intuitions. Since we therefore evidently do not all share the same taste for what is beautiful and what are supposed to be the relevant non-aesthetic features (especially, their weighting) in which the beautiful is grounded, it is only to be expected that we will not agree about which theory is the best candidate to choose. So, the claim that theories are to be evaluated according to their aesthetic properties does not help us in finding an *agreement* when selecting one candidate as the winner in the game, but it does help us in seeing *why* there is no such agreement. It helps us to see that there are different equally good theories available, and that there is no meta-theoretical criterion such that it would clearly have one particular theory defeat all others. Furthermore, it makes us realise that, individually, we are still justified in selecting one theory as the most preferable according to its aesthetic properties, which in turn depend both on its non-aesthetic properties, on the context of their creation, and on our personal taste, whose role it is to determine which of the non-aesthetic features (according to

different weightings we might also importantly apply) are to give rise to the alleged beauty of the theory as a whole.

In a nutshell, here is then how the whole evaluative process of theories might be understood. First, the evaluator carefully examines a theory's theoretical virtues, such as explanatory power, simplicity, parsimony, compatibility with intuitions, etc. in order to see which ones the theory possesses and to what degree. To this end, the evaluator must be a trained and competent philosopher. Evaluations of this kind are sometimes an objective matter, a matter on which different philosophers can agree, at least to some extent. For instance, it will usually not be very difficult to agree that nominalism fares better with respect to qualitative parsimony than the theory of Platonic universals. But, of course, not all such evaluations are as easily decided: for instance, it is very much debated whether endurantism or perdurantism is better with respect to being compatible with our intuitions about personal identity. Thus, in many cases, these "first-order" evaluations are not something philosophers can easily agree on, including their status as being 'objective'. Indeed, already at this first stage, the evaluator will appeal to her personal preference, which is a kind of "philosophical taste", for such-and-such a way of seeing things.

Then we see the second stage, where *one* philosopher considers several competing theories that she has finished evaluating (that is, she has gone through the first stage on her own and has arrived at a firm opinion, for example, in that she believes that endurantism accommodates our intuitions about personal identity better than perdurantism). At this second stage, she has to choose, say, between endurantism and perdurantism. Suppose she thinks that endurantism is better with respect to our common sense intuitions and, in addition, is a more complex and intricate view, while perdurantism is a more revisionary, bolder, and structurally simpler view (as part of a more comprehensive assessment, she would voice an opinion on all the evaluative criteria mentioned above, of course). Her philosophical taste will here again play a role and tell her to go for one view rather than another, based on her aesthetic preference, say, for simplicity rather than complexity (and so, for perdurantism rather than endurantism). It is clear, then, that the evaluator's taste plays a role from the beginning to the very end of the evaluative process.

The Three Faces of Anti-realism

§11. In a sense, I already said everything I wanted to say in this book. I wanted to make some points about metaphysical equivalence, about primitive problem-solvers and their importance, and about the fact that we find ourselves in a very difficult situation when it comes to choosing between competing metaphysical theories. I also wanted to articulate and defend the idea that metaphysical theories are (or aren't) beautiful and that this is what often drives us to prefer one theory to another—thus making our meta-theoretical situation a little less difficult. I said everything I wanted to say, but something important remains. It's something that is not an unavoidable consequence of the claims I have defended, but it's something that—to my own

mind—stems from them. So, let me say more about a very general picture of the nature of metaphysics that I am inclined to take on board.

In his influential series of lectures entitled "The Many Faces of Realism", as well as in his other rich and significant work, Hilary Putnam laid down some milestones which shaped the debate between realists and anti-realists for the next decades, and which constitute the background net of ideas any contemporary discussion cannot do without. This being said, the two opposing doctrines—metaphysical realism and anti-realism—have since taken various forms, and have evolved in very different directions. To my mind, in this book, I have provided three reasons in favour of a kind of anti-realism—three faces of antirealism. They are: (i) considerations about theory choice (Chaps. 5 and 7), (ii) cases of metaphysical equivalence (Chaps. 1, 2, and 4), and (iii) the divorce of phenomenology/experience and metaphysics (Chap. 6). Let me discuss them in turn, in the light of the discussion from this chapter.

Indeed, a natural and useful way of understanding the results from the discussion above is to endorse the view that we should abandon the idea that metaphysical theories are true/false. Instead, I want to suggest that they are (i) useful models (and there can be more than one model for solving a given metaphysical problem[5]), and (ii) something like works of art that we appreciate for their beauty. Thus, embracing the claim that aesthetic features of philosophical theories are good guides when it comes to theory evaluation and theory choice, has a natural consequence that might not be of everybody's meta-philosophical taste: a kind of anti-realism, at least when it comes to the metaphysical theories on which I focused in this book. The worry one can have here can be simply put thus: even if we agree that theories possess aesthetic properties such as "being beautiful", why should beauty of a theory be a good criterion for its evaluation, since there seems to be no good reason to think that beauty is *truth-conducive*[6]? Indeed, if we aim at metaphysical *truth*, why should we think that beauty will lead us there? A view which would claim that beauty *is* truth-conducive, could still provide an explanation of why different philosophers hold different views in a sense relevant to our discussion, but with the important difference that the claim here would be that only one of us is right—only one philosopher's choice leads to an adequate description of the way the world is. If beauty is truth-conducive then only one philosopher's aesthetic evaluations will guide her to metaphysical truth, the others being, more or less, mislead. This is actually not completely implausible since, remember, aesthetic properties of metaphysical theories are grounded in their non-aesthetic features, and the claim would here then mean that the good philosopher

[5]Compare to Rescher (1978, p. 220): "Philosophical argumentation is accordingly *nonpreemptive*: the existence of one cogent resolution of an issue does not block the prospect of an equally cogent basis for its alternatives; by positive argumentation an excellent case can be built up in substantiation of each of several mutually incompatible theses. […] In philosophy, supportive argumentation is never alternative precluding. […] Every philosophical problem thus admits of a variety of mutually conflicting solutions on whose behalf an impressively cogent case can be made out."

[6]Compare to Todd (2008) whose concern is physics and not metaphysics, and who claims that aesthetic properties do not help in the matter of *empirical adequacy* of the theory.

is the one who has good taste with respect to those features and who is sensitive to the theory's beauty in the most relevant way, very much like a good art critic is capable of providing the best evaluations of some pieces of artworks by having her sensibility more finely tuned than her fellow colleagues.

To my mind, though, the view that beauty is truth-conducive is under-motivated. Naively asked: are we here invited to think that "the world is beautiful" in some sort of objective way and that beautiful theories are more likely to be true—that they represent the world correctly? Why suppose that the world is such?

It seems, then, a better strategy, in order to face the worry above, to *accept* that beauty is *not* truth-conducive, but claim that it still is the best guide when it comes to theory-choice. Behind this strategy lies a view of philosophy in general and metaphysics in particular as being an enterprise whose task is to analyse, organize and systematize our concepts, such as the concept of a material object or the concept of a property, in order to provide a better understanding of them and to show how they are able to explain some phenomena we encounter (such as attribute agreement, change over time,…) and to explore how these concepts are related to one another. A different way to put this is to insist that metaphysics does *not* say how the *world* is, it says what our concepts are like. (I want to stress that this claim is to be distinguished from the claim that metaphysics is conceptual analysis—a claim I of course do not (want to, need to) endorse[7]). This picture of the way metaphysics works goes then as follows:

(i) we start with the concepts we have (like the concept of an ordinary material object), where such concepts can, in principle, both be given to us a priori or from experience (more on this below)
(ii) we want to better understand the concepts we have, and understand the connections and relations between them
(iii) to this end, we may introduce new theoretical concepts (like a substratum, or a relation of compresence; often, these will be primitive problem-solvers)
(iv) this gives rise to a theory that is an organization and systematization of our concepts including the new theoretical ones (note that while doing this it may so happen that we are forced to revise or abandon some of the concepts we started with)

If this is correct, that is, if metaphysical theories are about our concepts rather than about the world, then when we evaluate metaphysical theories, we should not ask "is this theory correctly and adequately describing the world?"—rather we should ask "is this theory a good/bad analysis and systematization of our concepts?" Thus, when facing a situation such as the case we have seen of the theory of universals, trope theory, and nominalism, we can, and we should, recognize that there are three *equally good* theories. Indeed, all three theories have their pro and con

[7]There are two central claims typically endorsed by friends of conceptual analysis which I reject (see Chap. 6): first, it is wrong to claim that metaphysics is a purely armchair a priori affair, and second, the weight defenders of conceptual analysis often put on the role our intuitions play in metaphysics is misguided.

arguments, all three can be evaluated in terms of their meta-theoretical non-aesthetic features, but still we are in a position where there is no objective consensus and no objective way to select one of them as being a clear winner of the game. Another way to put this is to say that metaphysical theories are theoretical and conceptual *models* of reality—where alternative, equally good, models can be available.[8]

But wait, one of the theories says that there are substrata, while the other two deny this claim—so surely, only one of them can be right! No, because what we do when we say that there is a substratum is not to say what there is in the world, rather, we introduce a new theoretical concept that allows to systematize, organize, and understand the concepts of material object and property in such a way that we have a satisfactory answer to the questions we started with. But, as the dialectical situation of the example of our three theories shows, this answer can also be given within a different framework of concepts that do not include a substratum but instead, say, a concept of resemblance or a concept of compresence. Thus, endorsing the view that metaphysics is about our concepts allows one to abandon the fantasy that there is *one* true theory about the world to be found, and gives one the possibility of acknowledging that there are *several equally good (even if conflicting) theories* around.

Bearing such a view in mind, we can now lighten our burden concerning the worry that there is no good reason why aesthetic properties of theories should be *truth-conducive*. Indeed, it can be acknowledged that beauty is *not* truth-conducive in the sense of 'finding out the theory which adequately describes the world', but this does not create a problem since if metaphysics is not about *describing the world* but about analyzing and systematizing our concepts, such as the concept of an object or a concept of a property, the need for any link between beauty and empirical adequacy or adequacy of a description does not arise. We can thus have a good answer to the question why different equally well informed and rational philosophers diverge in their choice of the best theory, while avoiding the worry from truth-conduciveness by simply denying the need for it.

Thus, the failure of the traditional meta-theoretical criteria to adjudicate between metaphysical theories (Chaps. 4–6) *suggests* such an anti-realist view of metaphysics. As an argument for these or other forms of anti-realism we can then point out that disagreement among equally rational and well-informed practitioners of metaphysics is best explained by the suggestion that aesthetic considerations are driving theory choice, and the fact that aesthetic considerations are driving theory

[8]Compare to Paul (2012, my italics) who seems to hold a similar view concerning this particular point, but restricts her claim only to a sub-class of metaphysical theories: "Metaphysical theories *exploring parts of the world that are in principle accessible to scientists* should be taken *as describing toy models* of the empirical facts, where such models represent ways the world might be, given the information we have to date. These models can be compared in terms of elegance, simplicity, empirical adequacy (to the extent that empirical facts are known) and consistency with contemporary science, *but should not be adopted as true*. […] Science and empirical discoveries will ultimately determine which, if any, of the toy models provided by metaphysicians should be given the status of a true theory of the world."

choice at least seems to suggest that metaphysics isn't aiming at truth about the world, in the sense of providing an adequate description of it. In short, here is the path to the kind of anti-realism that stems from our discussion about theory choice: first, we are not in a position to be able to choose between competing metaphysical theories, such as the trope theory, the theory of universals, or resemblance nominalism, by appealing to their pros and cons; second, we are not able to make a choice by appealing to standard meta-theoretical evaluative criteria; third, this provides ground for the idea that these are different but equally good theories; fourth, their aesthetic properties do help us to make a choice, but are not truth-conducive in the sense discussed above; thus, our metaphysical theories are best seen as organizations and systematizations of our concepts (*en passant*, think of my discussion of Thales in §8 above) rather than as descriptions of the world—the idea that there is one good, true, privileged theory among our candidates is a myth. (Ted Sider's discussion of the "red-blue universe" is a very good example of the exact opposite of what I want to convey here (see Sider 2011, Chap 1, §1.1–3)).

A similar path is suggested by our discussion from Part I of metaphysical equivalence. Indeed, as we have seen, the fact that there are cases of equivalent theories (such as the bundle theory and the substratum theory, or relationism and substantivalism), or cases of partially equivalent theories (such as, endurantism and perdurantism), provides grounds again for the claim that they are purely alternative formulations—equally good, even if different.

Let us be more precise about the relationship between my metaontological claims from Part I and the claims I am making here about the nature of metaphysics in general. In Part I, I have claimed that some theories are metaphysically equivalent, and that these are examples of merely verbal disputes. But I have also claimed that to get to this result, we need to go through a careful analysis of particular debates, and we need to see how the theories work in detail. Thus, if we find cases of disagreement—that is, of *non*-equivalent theories—we have then cases of substantive, non-verbal, disputes. Doesn't this conflict with the anti-realist view I am suggesting now that there is nothing substantive about metaphysics in general, in the sense that metaphysics does not tell us how the world is but rather tells us what our concepts are like? Well, if by "substantive" one means "ontologically deep", then there would be a tension between my earlier claims and the claims I am making now. But when I speak of metaphysical equivalence—and *non*-equivalence—in Part I, I do not have (to have) such a strongly realist understanding of "substantive" in mind. Rather, by "substantive" I mean that the disagreement is not generated by some trivial and merely verbal misunderstanding that could be resolved by showing that the two parties are using terms which belong to different linguistic frameworks. Thus, there is room for genuine and substantive disagreement, but this disagreement takes place in disputes that are about our concepts (and about primitives).

Finally, in Chap. 6, we have seen a different—almost Kantian—reason to think better of a kind of anti-realist view, namely, the divorce of phenomenology/experience and metaphysics. It is useful here to compare our situation to the debate about scientific (anti-)realism and the 'underdetermination of theory by data' argument. Often traced back to Duhem (1906), the idea is that different scientific

theories are consistent with the same set of data—thus, theory choice is underdetermined by the data. In short, given the data we have, we cannot say which of the competing theories is 'the correct one'. To put it differently, the observational data are neutral to with respect to the different theories, all compatible with it. As we have seen, in many cases, the situation we find ourselves in when it comes to metaphysical theories is very similar. Indeed, we have seen that our experience of the world is one thing, and metaphysical reality is another (think again, for instance, of the case of the A-theory and B-theory of time, and of how our temporal experience claims are simply orthogonal to the metaphysical issue). As a consequence, the choice of one metaphysical theory as being 'the correct one' is here also underdetermined by the data—the data being our experience and our intuitions. Indeed, concerning intuitions, we have seen that often our so-called philosophical intuitions are no more than constructs out of the contingent way we experience the world, and that they are not trustworthy guides when it comes to metaphysical claims. This left us in an uncomfortable situation since there is then the risk that at the end of the day, when we realize this, and when at the same time we recognize that in many domains of metaphysics intuitions play a central role in both theory construction and theory evaluation, we find ourselves in an impasse since more often than not we just don't have at our disposal any other data than these so-called intuitions. If we cannot build our metaphysical theories on such unreliable data, on what are we going to build them?

As before, even though it is not a direct and unavoidable consequence of the discussion from Chap. 6, I think that one possible and very natural reaction to these worries could be a form of anti-realism. We could, here again, simply accept that the best we can do when doing metaphysics is to understand the world as it is given to us (perceptually and conceptually), and not as it is. Indeed, this 'third face' of anti-realism fits well with what I just said above concerning theory choice and metaphysical equivalence: a natural way of understanding the divorce of phenomenology/experience and metaphysics is to insist that metaphysics does not tell us how the world is, but that it tells us what our concepts are like (again, this claim is to be distinguished from the claim that metaphysics is conceptual analysis). According to this picture of the way metaphysics works, we start with the concepts—and, yes, intuitions—we have (like the concept of an ordinary material persisting object, and experiential and intuitive data associated with it), where such concepts can, in principle, both be given to us a priori or from experience. Intuitions, as we have seen, seem to come often from the phenomenal character of our experience. These are the only data we can possibly have when it comes to many metaphysical issues (on which physics and other sciences are simply silent). As we advance in the metaphysical enterprise, we want to better understand these concepts (and percepts, and intuitions), and we want to understand the connections and relations between them. To this end, we may use empirical science. Physics can prove useful when it comes to some metaphysical debates (perhaps, the presentism vs. eternalism controversy), but not in many other cases (say, the tropes vs. universals debate, statues and lumps, eliminativism, mereological composition, personal identity, and many other). Experimental psychology and cognitive science, as we have seen in Chap. 6,

can prove to be extremely useful in many cases, mostly to assess the value of the intuitive and experiential data we started with. Finally, as metaphysicians, we may want to (need to) introduce new theoretical concepts (say, bare particulars or relations of compresence, or temporal counterpart relations, etc.). In the end, this procedure gives rise to metaphysical theories which are organizations and systematizations of our concepts, the surviving old ones as well as the new ones. And, crucially, *that's it*. That's as far as we can go.

References

Adams, R.M. 1974. Theories of actuality. *Noûs* 8: 211–231.

Armstrong, D.M. 1978. *Nominalism and realism*. Cambridge: Cambridge University Press.

Armstrong, D.M. 1997. *A world of states of affairs*. Cambridge: Cambridge University Press.

Audi, P. forthcoming. A clarification and defense of the notion of grounding. In *Grounding and explanation*, eds. F. Correia and B. Schnieder. Cambridge: Cambridge University Press.

Bealer, G. 1987. The philosophical limits of scientific essentialism. *Philosophical Perspectives* 1: 289–365.

Bealer, G. 1998. Intuition and the autonomy of philosophy. In *Rethinking intuition*, ed. M. DePaul, and W. Ramsey. Oxford: Rowman & Littlefield Publishers.

Beedle, A. 1996. Modal fatalism. *The Philosophical Quarterly* 46: 488–495.

Bennett, K. 2004. Spatio-temporal coincidence and the grounding problem. *Philosophical Studies* 118(3): 339–371.

Bennett, K. 2008. Composition, colocation, and metaontology. In *Metametaphysics*, ed. David Manley, David Chalmers, and Ryan Wasserman. Oxford: Oxford University Press.

Bennett, K., and B. MacLaughlin. 2005. Supervenience. In *Stanford Encyclopedia of philosophy*.

Benovsky, J. 2006. *Persistence through time and across possible worlds*. Ontos Verlag.

Bigelow, J. 1996. Presentism and properties. In *Philosophical Perspectives*, vol. 10, ed. J. Tomberlin. Malden: Blackwell.

Black, M. 1952. The identity of indiscernibles. *Mind* 61: 153–164.

Bricker, P. 2006. The relation between the general and the particular: Entailment vs. supervenience. In *Oxford studies in metaphysics*, vol. 2, ed. D. Zimmerman, 251–287. Oxford: Oxford University Predd.

Brogaard, B. 2000. Presentist four-dimensionalism. *The Monist* 83: 341–354.

Burke, M. 1994. Preserving the principle of one object to a place: A novel account of the relations among objects, sorts, sortals, and persistence conditions. *Philosophy and Phenomenological Research* 54(3): 591–624.

Callender, C. 2000. Shedding light on time. *Philosophy of Science* 67: S587–S599. (Proceedings).

Campbell, K. 1981. The metaphysics of abstract particulars. Reprinted in Macdonald, C. and C. Laurence. 1998. *Contemporary readings in the foundations of metaphysics*, 329–350. Oxford: Basil Blackwell.

Chalmers, D. 1996. *The conscious mind: in search of a fundamental theory*. Oxford: Oxford University Press.

Chalmers, D. 2008. Ontological anti-realism. In *Metametaphysics*, ed. David Manley, David Chalmers, and Ryan Wasserman. Oxford: Oxford University Press.

Chisholm, R. 1990a. Events without times: An essay on ontology. *Noûs* 24: 413–428.

Chisholm, R. 1990b. Referring to things that no longer exist. In *Philosophical perspectives*, vol. 4, ed. J. Tomberlin, 545–556.

Correia, F. 2008. Ontological dependence. *Philosophy Compass* 3(5): 1013–1032.

Craig, W.L. 2000. *The tensed theory of time*. Dordrecht: Kluwer.

J. Benovsky, *Meta-metaphysics*, Synthese Library 374,
DOI 10.1007/978-3-319-25334-3

Derkse, W. 1992. *On simplicity and elegance*. Delft: Eburon.
de Rosset, L. 2010. Getting priority straight. *Philosophical Studies* 149: 1.
de Witt, B., and N. Graham. 1973. *The many-worlds interpretation of quantum mechanics*. Princeton: Princeton University Press.
Dirac, P.A.M. 1980. The excellence of Einstein's theory of gravitation. In *Einstein: The first hundred years*, ed. M. Goldsmith, A. Mackay, and J. Woundhuysen. Oxford: Pergamon Press.
Dorato, M. 2006. The irrelevance of the Presentist/Eternalist debate for the ontology of Minkowski spacetime. In *Philosophy and foundations of physics. The ontology of spacetime*, ed. D. Dieks. Amsterdam: Elsevier.
Duhem, P.M.M. 1906. *The aim and structure of physical theory*. Wiener (tr.). Princeton: Princeton University Press.
Ehring, D. 2001. Temporal parts and bundle theory. *Philosophical Studies* 104: 163–168.
Everett, H. 1957. Relative state formulation of quantum mechanics. *Reviews of modern physics* 29: 454–461.
Fine, K. 1995. Ontological dependence. In *Proceedings of the Aristotelian Society*, vol. 95, 269–290.
Forbes, G. 1993. Time, events, and modality. In *The philosophy of time*, ed. R. Le Poidevin, and M. MacBeath. Oxford: Oxford University Press.
Haslanger, S. 2003. Persistence through time. Reprinted in Loux, M.J. and D.W. Zimmerman, eds. 2003.
Hawley, K. 2006. Science as a guide to metaphysics? *Synthese* 149: 451–470.
Hawthorne, J. 1995. The bundle theory of substance and the identity of indiscernibles. *Analysis* 55 (3): 191–196.
Hawthorne, J., and J. Cover. 1998. A world of universals. *Philosophical Studies* 91: 205–219.
Heller, M. 1990. *The ontology of physical objects: Four-dimensional Hunks of Matter*. Cambridge: Cambridge University Press.
Heller, M. 2003. The immorality of modal realism, or: How I learned to stop worrying and let the children drown. *Philosophical Studies* 114.
Heller, M. 2008. The donkey problem. *Philosophical Studies* 140.
Hirsch, E. 2005. Physical-object ontology, verbal disputes, and common sense. *Philosophy and Phenomenological Research* 70: 67–97.
Hirsch, E. 2007. Ontological arguments: Interpretive charity and quantifier variance. In *Contemporary debates in metaphysics*, ed. John Hawthorne, Theodore Sider, and Dean Zimmerman. Amsterdam: Blackwell.
Hirsch, E. 2008. Ontology and alternative languages. In *Metametaphysics*, ed. David Chalmers, David Manley, and Ryan Wasserman. Oxford: Oxford University Press.
Hume, D. 1985. Of the standard of taste. In *Essays: Moral, political, and literary*, ed. E. Miller. Indianapolis: LibertyClassics.
Jackson, F. 1994. Armchair metaphysics. In *Mind, method, and conditionals: Selected essays*, ed. F. Jackson, 154–176. London: Routledge.
Jackson, F. 1998. *From metaphysics to ethics*. Oxford: Oxford University Press.
Johnston, M. 1987. Is there a problem about persistence ? *The Aristotelian Society* 61: 107–135.
Kelly, T. 2005. The Epistemic Significance of Disagreement. In *Oxford studies in epistemology*, vol. 1, ed. John Hawthorne, and Tamar Szabo Gendler, 167–196. Oxford: Oxford University Press.
Kim, J. 1984. Concepts of supervenience. In *Philosophy and phenomenological research*, vol. XLV, no. 2.
Kim, J. 2006. *Philosophy of mind*, 2nd edn. Westview Press.
Kuhn, T.S. 1973. Objectivity, value judgment, and theory choice. In *The essential tension: Selected studies in scientific tradition and change,* 1977. Chicago: University of Chicago Press.
Ladyman, J., and D. Ross. 2007. *Every thing must go: Metaphysics naturalised*. Oxford: Oxford University Press.
Le Poidevin, R. 1991. *Change, cause, and contradiction*. London: Macmillan.

Le Poidevin, R. 2009. The experience and perception of time. In *Stanford Encyclopaedia of philosophy* (online).
Levinson, J. 1984. Aesthetic supervenience. *Southern Journal of Philosophy* 22(Supplement): 93–110.
Lewis, D. 1973. *Counterfactuals*. Oxford: Basil Blackwell.
Lewis, D. 1976. The paradoxes of time travel. *American Philosophical Quarterly* 13: 145–152.
Lewis, D. 1983. New work for a theory of universals. *Australasian Journal of Philosophy* 61.
Lewis, D. 1986. *On the plurality of worlds*. Blackwell.
Lewis, D. 1994. Reduction of mind. In *Papers in metaphysics and epistemology*, ed. D. Lewis. 1999. Cambridge: Cambridge University Press.
Lewis, D. 2002. Tensing the copula. *Mind* 111: 1–13.
Lombard, L.B. 1999. On the alleged incompatibility of presentism and temporal parts. *Philosophia* 27(1–2): 253–260.
Lowe, E.J. 1987. Lewis on perdurance vs endurance. *Analysis* 47: 152–154.
Lowe, E.J. 1988. The problems of intrinsic change: Rejoinder to Lewis. *Analysis* 48: 72–77.
Loux, M.J. 1997. Beyond substrata and bundles. Reprinted in *Contemporary readings in the foundations of metaphysics,* eds. S. Macdonald and C. Laurence. 1998. Oxford: Basil Blackwell.
Loux, M.J. (ed.). 1998. *Metaphysics: A contemporary introduction*. London: Routledge.
Loux, M.J. (ed.). 2001. *Metaphysics contemporary readings*. London: Routledge.
Loux, M.J., and D. Zimmerman (eds.). 2003. *The Oxford handbook of metaphysics*. Oxford: Oxford University Press.
Lowe, E.J. 1998. *The possibility of metaphysics*. Oxford: At the Clarendon Press.
Machery, E., R. Mallon, S. Nichols, and S. Stich. 2004. Semantics, cross-cultural style. *Cognition* 92(3): B1–B12.
Mackinnon, J.E. 2001. Aesthetic supervenience: For and against. *British Journal of Aesthetics* 41 (1).
McLaughlin, B.P. 1995. Varieties of supervenience. In *Supervenience: New essays*, ed. E. Savellos, and U. Yalcin. Cambridge: Cambridge University Press.
MacLaurin, J., and H. Dyke. 2012. What is analytic metaphysics for? *Australasian Journal of Philosophy* 90(2): 291–306.
Manley, D. 2002. Properties and resemblance classes. *Noûs* 36: 75–96.
Markosian, N. 2004. A defense of presentism. In *Oxford studies in metaphysics*, vol. 1, ed. D. Zimmerman. Oxford: Oxford University Press.
Martin, C.B. 1980. Substance substantiated. *Australasian Journal of Philosophy* 58: 3–10.
Maudlin, T. 2007. *The metaphysics within physics*. Oxford: Oxford University Press.
McAllister, J. 1999. *Beauty & revolution in science*. Cornell University Press.
McCall, S., and E.J. Lowe. 2003. 3D/4D equivalence, the twins paradox and absolute time. *Analysis* 63(278): 114–123.
McTaggart, J.M.E. 1927. *The nature of existence*, vol. 2. Cambridge: Cambridge University Press.
Melia, J. 2005. Truthmaking without truthmakers. In *Truthmakers: The contemporary debate*, ed. H. Beebee, and J. Dodd. Oxford: Oxford University Press.
Mellor, D.H. 1981. *Real Time*. Cambridge: Cambridge University Press.
Mellor, D.H. 1998. *Real Time II*. London: Routledge.
Merricks, T. 1995. On the incompatibility of enduring and perduring entities. *Mind* 104: 523–531.
Merricks, T. 2001. *Objects and Persons*. Oxford University Press.
Merricks, T. 2007. *Truth and ontology*. Oxford: Oxford University Press.
Miller, K. 2005a. The metaphysical equivalence of three and four dimensionalism. *Erkenntnis* 62: 91–117.
Miller, K. 2005b. What is metaphysical equivalence? *Philosophical Papers* 34(1): 44–74.
Mellor, D.H. 1998. *Real Time II*. London: Routledge.
Newton, I. 1687. *Philosophiæ Naturalis Principia Mathematica*.

Nolan, D. 1997. Quantitative parsimony. *The British Journal for the Philosophy of Science* 48: 329–343.
Oaklander, N. 1991. A defense of the new tenseless theory of time. *Philosophical Quarterly* 41: 26–38.
Parsons, J. 2004. Review of 'Four-dimensionalism' by Theodore Sider. *Philosophical Quarterly* 54.
Paul, L.A. 2002. Logical parts. *Noûs* 36: 578–596.
Paul, L.A. 2006. Coincidence as overlap. *Noûs* 40: 623–659.
Paul, L.A. 2010. Temporal experience. *Journal of Philosophy* 107(7).
Paul, L.A. 2012. Metaphysics as modeling: The Handmaiden's tale. *Philosophical Studies*.
Paul, L.A. forthcoming_a. Identity, indiscernibility and objects. In eds., Paul Loux and van Inwagen. forthcoming. Oxford University Press.
Paul, L.A. forthcoming_b. A one category ontology. In eds., John Adorno Keller. forthcoming. Oxford: Oxford University Press.
Pettit, P. 1987. The possibility of aesthetic realism. In *Pleasure, preference and value*, ed. F. Schaper. Cambridge: Cambridge University Press.
Prior, A.N. 1970. The notion of the present. *Studium Generale* 23: 245–248.
Prior, A. 1996. Some free thinking about time. In *Logic and reality: Essays on the legacy of arthur prior*, ed. J. Copeland. Oxford: Oxford Clarendon Press.
Prior, A.N. 2003a. *Papers on Time and Tense*, New ed. Per Hasle: Peter Ohrstrom.
Prior, A.N. 2003b. Changes in events and changes in things. in Prior, 2003a, 7–19.
Quine, W.V. 1948. On what there is. *Review of Metaphysics* 2: 21–38.
Quine, W.V.O. 1960. *Word and object*. Cambridge: M.I.T. Press.
Rescher, N. 1978. Philosophical disagreement. An essay towards orientational pluralism in metaphilosophy. *Review of Metaphysics* 32(2): 215–217.
Rodriguez-Pereyra, G. 2002. *Resemblance nominalism, a solution to the problem of universals*. Oxford: Oxford University Press.
Ruben, D. 1990. *Explaining explanation*. Routledge.
Russell, B. 1912. *Problems of philosophy*. Oxford: Oxford University Press.
Schaffer, J. 2009. On what grounds what. In *Metametaphysics*, ed. David Chalmers, David Manley, and Ryan Wasserman. Oxford: Oxford University Press.
Schaffer, J. manuscript. Grounding as the primitive concept of metaphysical structure.
Scholl, B.J. 2007. Object persistence in philosophy and psychology. *Mind and Language* 22(5): 563–591.
Scruton, R. 1974. *Art and imagination: A study in the philosophy of mind*. St. Augustine's Press.
Sibley, F. 1959. Aesthetic concepts. *Philosophical Review* 68: 421–450.
Sidelle, A. 2002. Is there a true metaphysics of material objects? *Philosophical Issues* 12.
Sider, T. 2000. The stage view and temporary intrinsics. *Analysis* 60: 84–88.
Sider, T. 2001a. *Four-dimensionalism*. Clarendon Press.
Sider, T. 2001b. Criteria of personal identity and the limits of conceptual analysis. *Philosophical Perspectives* 15: 189–209.
Sider, T. 2001c. Maximality and intrinsic properties. *Philosophy and Phenomenological Research* 63: 357–364.
Sider, T. 2006. Bare particulars. *Philosophical Perspectives* 20: 387–397.
Sider, T. 2007. Which disputes are substantive (unpublished).
Sider, T. 2008. Ontological realism. In *Metametaphysics*, ed. David Chalmers, David Manley, and Ryan Wasserman. Oxford: Oxford University Press.
Sider, T. 2011. *Writing the book of the world*. Oxford: Oxford University Press.
Sider, T. 2013. Against parthood. In *Oxford studies in metaphysics*, vol. 8, eds. Karen Bennett and Dean W. Zimmerman, 237–293. Oxford: Oxford University Press.
Skow, B. 2009. Relativity and the moving spotlight. *The Journal of Philosophy* 12: 666–678.
Simons, P. 1994. Particulars in particular clothing: Three trope theories of substance. *Philosophy and Phenomenological Research* 54: 553–575.

Simons, P. 2000a. How to exist at a time when you have no temporal parts? *The Monist* 83: 419–436.
Simons, P. 2000b. Continuants and occurrents. *The Aristotelian Society* 74: 78–101.
Shoemaker, S. 1969. Time without change. *Journal of Philosophy* 66: 363–381.
Shoemaker, S. 1997. Self and substance. *Philosophical Perspectives* 11.
Sober, E. 1981. The principle of parsimony. *The British Journal for the Philosophy of Science* 32: 145–156.
Strawson, G. 1997. The self. *Journal of Consciousness Studies* 4: 405–428.
Swain, S., J. Alexander, and J.M. Weinberg. 2008. The instability of philosophical intuitions: Running hot and cold on truetemp. *Philosophy and Phenomenological Research* 76(1): 138–155.
Todd, C. 2008. Unmasking the truth beneath the beauty: Why the supposed aesthetic judgements made in science may not be aesthetic at all. *International Studies in the Philosophy of Science* 22.
Vallicella, W. 1997. Bundles and indiscernibility: A reply to O'Leary-Hawthorne. *Analysis* 57(1): 91–94.
Van Cleve, J. 1985. Three versions of the bundle theory. Reprinted in Loux 2001.
Van Inwagen, P. 1981. The doctrine of arbitrary undetached parts. Reprinted in Van Inwagen 2001.
Van Inwagen, P. 1985. Plantinga on trans-world identity. Reprinted in Van Inwagen 2001.
van Inwagen, P. 2001. *Ontology, identity and modality*. Cambridge: Cambridge University Press.
Walton, K. 1970. Categories of art. *Philosophical Review* 79.
Walton, K. 2008. *Marvelous images: On values and the arts*. Oxford: Oxford University Press.
Weinberg, J.M., C. Gonnerman, C. Buckner, and J. Alexander. 2010. Are philosophers expert intuiters? *Philosophical Psychology* 23(3): 331–355.
Wiggins, D. 1968. On being in the same place at the same time. *Philosophical Review* 77(1): 90–95.
Williams, D.C. 1951. The myth of passage. *Journal of Philosophy* 48(15): 457–472.
Williams, D.C. 1953. On the elements of being. *The Review of Metaphysics* 7: 171–192.
Yablo, S. 2008. Must existence questions have answers. In *Metametaphysics*, ed. David Chalmers, David Manley, and Ryan Wasserman. Oxford: Oxford University Press.
Zangwill, N. 1994. Supervenience unthwarted: Rejoinder to wicks. *Journal of Aesthetics and Art Criticism* 52.
Zangwill, N. 2001. *The metaphysics of beauty*. Cornell University Press.
Zangwill, N. 2003. Beauty. In *Oxford companion to aesthetics*, ed. Jerrold Levinson. Oxford: Oxford University Press.
Zimmerman, D. 1996. Persistence and presentism. *Philosophical Papers* 25: 115–126.
Zimmerman, D. 1997. Chisholm and the essences of events. In *The philosophy of Roderick M. Chisholm*, ed. Lewis Hahn, 73–100. Peru: Open Court.
Zimmerman, D. 1998. Temporary intrinsics and presentism. In *Metaphysics: The big questions*, eds. P. Van Inwagen and D.W. Zimmerman. Amsterdam: Blackwell Publishers.