EMPLOYING RESOURCES FROM SOCIAL ONTOLOGY, CONCEPTUAL ENGINEERING, AND CONCEPTUAL ETHICS TO ASSESS THE SOCIAL, AGENT-DEPENDENT FEATURES OF OBJECTS

A Dissertation

by

JARED SCOTT OLIPHINT

Submitted to the Graduate and Professional School of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Christopher Menzel
Co-Chair of Committee,	José Bermúdez
Committee Members,	Nathan Howard
	Robert Garcia
	Harold Boas
Head of Department,	Theodore George

December 2022

Major Subject: Philosophy

ABSTRACT

In what ways do metaphysical theories, language about objects in the world, and values we assign to those objects depend on social agents? I introduce a metaphilosophical model for how some of the characteristics of entities depend on social agents, particularly the characteristics involving language and values assigned to entities by social agents. In the first chapter I describe and summarize many of the positions I discuss in chapters two and three, including object theories, a model of social kinds, the project of conceptual engineering and conceptual ethics, and the metaphysics of words. Beginning in chapter two, I argue that 1) a two-dimensional model of social kinds is superior to the type-token model for explaining the conditions for a given type or kind, and 2) that the model's superiority is further shown when the model incorporates some of the features from the project of conceptual engineering and conceptual ethics. Because I will be talking about *objects* with social features, I need to explain what I mean by "object", given the differing object theories. And because many social objects like groups and money have value, there may be normative, conceptual ethics considerations, based on those objects' value, that affect the conditions for what is preferred to count as an instance of OBJECT under those theories. We have at least *prima facie* motivation for thinking that we should prefer an object

theory that is most consistent with affirming that there are such valuable objects, so I consider how each object theory accommodates that preference.

But that raises the question of which objects are social. So I then argue in chapter three that characterizing objects as social in *degrees* rather than social or non-social *simpliciter* gives us the right characterization of objects that count as social. Many objects we find in the natural world that we might not ordinarily characterize as social — for example, artifacts that non-human animals use — nonetheless seem to have at least some social features. Despite what might be a *prima facie* motivation to characterize social objects as a completely separate class from non-social objects, I show that a disjoint division between social objects and non-social objects delivers several counterintuitive results. So I challenge such a characterization and then offer my own alternative view where objects are social to some degree. Finally, I then explain some of the puzzling features of one of the most significant kinds of objects in the social world: *words*. As parts of language, words as social objects function as the primary representational devices that social agents use to assign concepts and values. I argue that the metaphysical features of words as representational devices depend in many ways on social agents, and that this fact that is often overlooked explains some of the puzzling features of words and other social objects more broadly. I end with some concluding remarks.

DEDICATION

This dissertation is dedicated to the members of my family—both immediate and extended—who have shown such support throughout graduate school. To Kyle, Lizzie, and Caroline Oliphint, and to the families of Joel and Kate Oliphint and Bonnie and Andrew Matthews. To my parents, Scott and Peggy Oliphint, who offered more encouragement than I can express. And most of all to my wife Jenny, without whom this work truly would not have been possible, and to our beautiful children: Senna, Kylah, Clara, Seren (born during my first year), and Kase (born during my last year).

ACKNOWLEDGEMENTS

Many thanks go to my committee, who should be credited with any strengths of the work and none of the weaknesses. To Dr. Boas's labors and helpful comments as the outside member from the mathematics department. To Dr. Bermúdez for providing needed criticisms that made this final product better. To Dr. Howard, who showed me how to structure and pitch a paper from my earliest (poor) attempts at publishing. I will forever be grateful for his intentional mentorship, his example of what it looks like to turn out an astonishing output of papers of the highest quality, and for being a reasonable, ethical voice during many category-five storms throughout the program. To Dr. Garcia, who drew me to consider the program here in the first place and who I'm honored to consider a friend. Many thanks are due for showing me what a cogent, coherent argument should look like, and for the needed holistic, life advice while navigating the never-ending potential perils in this academic industry. And to Dr. Menzel, who was tasked with the burden of painstakingly showing me why every written word should serve the purpose of maximal clarity and precision. It was never lost on me that my talents fell far short of what should have been a more enjoyable advisory role, but I will continually be thankful to him for showing me what a well-crafted paper should look like.

I owe thanks to dozens, if not hundreds, of people outside the committee, and I'm grateful to have a space here that is allowed to be more warm-blooded than the rest of the project. To the many people here at A&M, including my many students, who guided and helped me through this process in all kinds of ways: to Michael Hand, who introduced me to the philosophy of language during my first two years. To Kristi Sweet, a model of rigor, professionalism, and sincerity I would have been privileged to learn more from had I gone down the Continental or history of philosophy road. To Kenny Easwaran, quite possibly the most brilliant mind I will ever meet. His guidance during my early days of the program was invaluable, especially through our directed readings, and whose care for graduate students was evident. To Ted George, whose wisdom and guidance proved immensely helpful from the first of the program to the last. Thanks to Denise Meda Calderon in my cohort, who from day one shared a relentless work ethic and refreshingly honest perspective on the challenges that come with graduate student life. To Victoria Green, co-laborer in an unreasonably shortened program length. To Brady DeHoust, fellow believer and encourager, with whom conversations involving the spiritual and theological were always cathartic. To Dr. Alexander Crist, buddy, who authentically lives out his Continental philosophy. Thanks are owed to him for being a social oasis through the many unrelenting trials during the torturously

long quarantine. And to Sean Conte, who has been alongside me in the graduate belly of the beast since day one. Many thanks to him for being an encouraging voice through those first seminars, TAing, and figuring out how to finish this thing.

I owe so many acknowledgements to friends outside the philosophy world. To Dr. Micah Green, for genuinely looking out for my well-being throughout the program. My experience would be far poorer without our many lunches and conversations. To Liz Steel, for showing the purest form of kindness to my daughters, and for being a model they will always look up to. To her father, Jon Anderson, along with Wade Coleman, who were true pastors and shepherds to our family over these years, showing encouragement and care that will continue with us. I am always grateful for Dr. Nathan Sasser's friendship since the days of standing in line for Summer Greek, but particularly during this time as he encouraged and advised through the many academic challenges familiar to him. To the many WTS friends I hope now to reconnect with, including Dave Garner, John and Rhonda Currie, Nate Shannon, and especially my lifelong friend, co-laborer, co-writer, and fellow Aggie Jonathan Brack. To Dan Kunkle, who believed in me as a high school punk when he had no reason to. A heartfelt thanks to Larry Westerveld; I am forever in his debt for his able, steadfast help through the ecclesial trenches of insanity.

Finally, there are no words that can express the depth of gratitude and love to my family, to whom this dissertation is dedicated. To Kyle and Lizzie Oliphint, who I wanted to see more but am thankful for the times we were able to spend, and for their kind generosity. And especially to their daughter Caroline, who filled this experience with joy for my family. My girls will have countless, rich memories of our time here because of her. To my siblings, Joel and Bonnie, and their families, I owe both thanks and apologies for being away for so long. Their prayers, support, and encouragement lifted us up so many times over these five years. To my parents, Scott and Peggy, who witnessed this dream emerge decades ago and now turn into a (*very* different) reality. Their generosity and care are truly supernatural, and their visits and conversations carried us through the many challenges. I will never be able to thank them enough, and now hope to make up for lost time. To my kids, who were an irreplaceable wellspring of joy every day: Senna, Kylah, Clara, Seren, and Kase. I owe them a lifetime of quality time and presence, and my hope is they take with them good memories while we made Aggieland our home. And to my heroic, lovely, and omnicompetent wife, Jenny, who went through the process and trials in almost every way I did; the completion of all this is due to her in so many ways. We have shared every level of every emotion numerous times over. Thanks ultimately to God for all these people and experiences, for granting me

perseverance, and for graciously seeing me through these trials despite my many shortcomings. SDG.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supervised by a dissertation committee consisting of Professor Menzel [advisor], Professor Bermúdez [co-advisor] of the Department of Philosophy, Professor Garcia of Baylor University of the Department of Philosophy, Professor Howard of the University of Toronto of the Department of Philosophy, and Professor Boas of the Department of Mathematics.

All other work conducted for the dissertation was completed by the student independently.

Funding Sources

Graduate study was supported by Texas A&M University and a summer dissertation fellowship from The College of Liberal Arts.

TABLE OF CONTENTS

Page

CHAPTER I

INTRODUCTION

My working assumption is that there are social objects; that is, objects whose features depend on social agents in some way. Can we better understand some features of objects that have typically been thought of as mind-independent (composition, kind membership) if we think of those features as dependent on social agents in some way? What distinguishes social objects from other kinds of objects? If objects used in language (like words) are social objects, in what ways do the features of some objects depend on our use of language, if at all? What normative (especially moral) considerations might there be if there are different ways for social agents to characterize a given object? These are some of the questions I address in this dissertation.

To address those questions, I employ insights from two relatively new fields that have devoted much attention to the social world: social ontology, and conceptual engineering and conceptual ethics (hereafter the latter two will together be called "CE"). Up until now, these fields have been examining many of the same subjects within the social world (e.g. race, gender) in isolation from one another, leading to missed opportunities for gleaning insights and results from each field. I offer a remedy for this isolation by combining CE with Brian Epstein's two-dimensional model of kinds, which includes grounding and

anchoring (described in the first chapter), into what I call the "2D-CE" model. Particularly distinctive of Epstein's model is his distinction between grounding and anchoring, that is, roughly, between facts about whether an object satisfies some relevant conditions and facts about whatever sets up those conditions. My view is that the utility of this distinction becomes even more apparent when the latter kinds of facts are recast within the context of conceptual engineering and conceptual ethics as conditions for satisfying a concept in question.

I show the utility of the combined model by arguing that in many cases it can be applied to objects that have typically been thought of as non-social objects. Though objects from the social world serve as motivators and entry points for work in social ontology and CE, I use the 2D-CE model to address how objects typically thought of as outside the scope of the social world may actually have social features, i.e. features that depend on agents in some way.[*] If the aims for this project are successful, not only will we have a useful, combined model from both fields that can be applied to the kinds of social objects mentioned above, but the combined model can capitalize on the breadth of application to other objects mentioned above and can extend to many sub-disciplines within philosophy.

---

[*] Some philosophers who are working in social ontology and CE explicitly state that their work is intended to be applied more broadly to objects that are typically considered to be beyond the social world.

With the above preliminary motivation in place, I begin in the next chapter by expositing much of the literature that chapters two and three rely on. The second chapter argues for the 2D-CE model and argues that permissivism is the object theory most compatible with social ontology. The third chapter attempts an answer to the question of what distinguishes social objects from other kinds of objects, and ends with applying the model to the social objects that are crucial for any project: *words*. In the conclusion I then briefly chart a path for further work.

I.1. References

Cappelen, Herman and Plunkett, David. (2020). "Introduction: A Guided Tour of Conceptual Engineering and Conceptual Ethics" in *Conceptual Engineering and Conceptual Ethics*. Alexis Burgess, Herman Cappelen, and David Plunkett, eds. Oxford University Press.

Dershowitz, Naomi. (2020). "Nihilism, But Not Necessarily." *Erkenntnis*, September. https://doi.org/10.1007/s10670-020-00311-7.

Epstein, Brian. (2019). Anchoring versus Grounding: Reply to Schaffer. *Philosophy and Phenomenological Research* 99 (3): 768–81. https://doi.org/10.1111/phpr.12644.

CHAPTER II

THE LITERATURE ON OBJECT THEORY, SOCIAL KINDS, AND CONCEPTUAL ENGINEERING

II.1. Introduction

This chapter reviews the literature that I will be drawing upon and interacting with in later chapters. Recall one of the primary questions of the project: can we better understand some features of objects that have typically been thought of as mind-independent (composition, kind membership) if we think of those features as dependent on social agents in some way? My aim for this project is to make a start at answering this question. The thesis of the project is two-fold. First, I argue that the 2D-CE model that combines the two-dimensional ("2D") model of kind membership with the projects of conceptual engineering and conceptual ethics ("CE") clarifies several issues within object theory and social ontology, particularly with application to the metaphysics of words. To argue for the superiority of the two-dimensional model of kind membership, I need to describe both that model and its foil, the type-token model. Second, I argue that the two-dimensional model is set up to provide an explanation of how it is the case that many *objects* (including *social* objects) can be unified under one *kind*. In support of this, I need to say a few things about object theory in this chapter so that in later chapters I can say more about what I mean by "object" and which

theory best supports an ontology of the social world. But I begin by introducing an important distinction regarding how we speak about many things as *one*, per Cotnoir (2013): *syntactic unity* and *semantic unity*. With that distinction in place, I introduce the most prominent object theories: nihilism, conservatism, and permissivism. These theories typically focus on "ordinary" composite objects, but, following current trends, I will assume that there are *social* objects exist as well. Accordingly, I briefly describe Passinsky's (2016) view for distinguishing social objects from non-social objects. I introduce Wetzel's (2009) type-token model as one account of how objects are unified, followed by Epstein's (2015) model for social kinds that I later argue is more helpful.

II.2. Syntactic Unity and Semantic Unity

In the context of formalizing composition as identity (CAI for short), Cotnoir (2013) presents the following distinctions:

> Consider the following two distinctions: *syntactically plural* vs *syntactically singular*, and *semantically plural* vs *semantically singular*.
>
> An English term is *syntactically plural* if it is morphologically plural, and *syntactically singular* if it is morphologically singular. The syntactic difference between singular and plural

> morphemes often takes the form of different endings (e.g. 'fungus'; 'fungi') and usually requires different verb conjugation (i.e. 'The dog barks.'; 'The dogs bark.')… Also, let a term be *semantically plural* if the referent of the term is more than one object, and let it be *semantically singular* if the referent is only one object. (Cotnoir 2013: 297)

In English, the syntactic construction does not always align with the semantic referent. Cotnoir gives a few helpful examples:

> Consider 'scissors' and 'pants'. Both have morphologically plural endings, conjugate plurally ('The scissors are sharp.'; *'The scissors is sharp.'), and yet putatively refer to a single thing. It may be noted that 'scissors' and 'pants' are linked to constructions like 'pair of scissors' and 'pair of pants', which are syntactically singular. But this simply reinforces the conclusion that, in English, we may use syntactically singular and plural constructions to refer to the very same thing(s).
>
> Likewise, we can find examples of syntactically singular terms that are semantically plural. Consider 'clothing', 'furniture', and 'jury'. These terms are morphologically singular, conjugate singularly ('The furniture is Federal.'; *'The furniture are

> Federal.'), and yet each of them refer to many things. (Cotnoir 2013: 297)

To get a sense of how Cotnoir's distinctions are relevant to my project, consider the following case from the literature on social ontology:

> [W]e submit that the Supreme Court is not one thing, but rather many things. The expression "the U.S. Supreme Court" is syntactically singular, yet semantically plural; and when used in the present tense, it plurally denotes the contemporaneous U.S. Supreme Court Justices…Syntactically singular group terms are disguised plurals. And syntactically plural group terms—like "the Supremes"—are undisguised plurals. (Horden and López de Sa 2020)

To deny that the Supreme Court is one thing, as in this quote, is to make an ontological claim that has semantic implications for what "the Supreme Court" refers to. Cotnoir more accurately captures the relation I want to consider here:

> Needless to say, whether a given term is semantically plural or singular may be controversial. It may even be controversial as to *whether there is any mind-independent fact of the matter* as to

> whether the referent of a term is many or one. (Cotnoir 2013: 297; emphasis mine)

Indeed, it may be that whether the referent of a term is one thing or many things depends in some way on agents, and is therefore socially dependent. The indeterminacy of whether a term picks out one thing or many things may depend in some way on agents, rendering the fact of the matter to be socially dependent. That is one of the primary metaphysical questions I will be addressing in this project. But before I take up that issue I will be looking at accounts in recent literature that have offered explanations for 1) what it is for many things to count as *one* object and 2) what it is for many things to be instances of *one* kind.

II.3. Object Theories

There is significant debate over what counts as an object, in particular over what counts as a *composite* object. Some philosophers have focused the debate on attempts to answer van Inwagen's (1990) *Special Composition Question* (*SCQ*), which asks under what conditions composition occurs. The main theories attempting to answer the SCQ fall into three broad categories: *nihilism*, *permissivism* (sometimes called *universal composition*), and *conservatism*. To simplify my description of these theories, I will assume *atomism*, the view that

there are objects in the universe that have no proper parts—in the literature on mereology, they are called *atoms, simples,* or sometimes *particles*.* The description of these three categories below will be drawn largely from Cotnoir and Varzi's (2021), particularly section 5.2 on existence and identity, and from Korman (2015).

II.3.1 Nihilism

Nihilism is one form of *eliminativism*, where eliminativism denies that composition occurs, except perhaps under certain restricted conditions. For example, the variety of eliminativism known as *organicism* (e.g. van Inwagen 1990; Merricks 2001) denies that composition occurs except in cases involving some biological (i.e. organic) objects. Nihilism denies even that—it is the view that "everything is mereologically atomic", i.e. that there are no objects with proper parts.† (Cotnoir and Varzi 2021: 181)

The benefit of nihilism is its *prima facie* simplicity: only atomic objects exist. This makes articulating nihilism as an ontological theory relatively easy. But the cost of the theory is high, because it denies the existence of almost all

* Precedent for such an assumption that is made for heuristic purposes can be found in Cotnoir and Varzi (2021: 181), to cite just one instance.
† There are several versions of eliminativism, however, like a form of eliminativism that believes things like tables exist but that they are not composite objects. There is also a form of eliminativism that even accepts universalism. See Korman (2015) section 2.1 for a helpful survey of other forms of eliminativism.

objects we ordinarily refer to and take to exist, such as tables, chairs, trees, and so forth (assuming those are composite objects). Insofar as the existence of a given social object depends on the existence of composite objects, the view would also deny the existence of many social objects like money and groups, because those social objects depend on composite objects like paper bills, coins, and persons (though *organicists* would affirm the existence of persons which, as I argue in chapter three, are social objects).

If ordinary objects like tables do not in any sense exist, much of our ordinary language turns out to be false or misleading. In response, nihilists argue that talk of tables could, in theory at least, be replaced by talk of "particles arranged table-wise"; similarly for other ordinary objects. But, as Cotnoir and Varzi (2021: 183) observe, many have pointed out that this move merely trades the Special Composition Question for the Special Arrangement Question: "given any composite-object sortal K, under what circumstances are there things arranged K-wise?" And, arguably, that question is no clearer and no less pressing than the SCQ.

II.3.2 Permissivism

My focus will be on a particular kind of permissivism, *universalism*, which Cotnoir and Varzi describe this way: "for any condition φ whatsoever, as long as it is satisfiable, there exists something composed of all and only the things that

satisfy that condition". (Cotnoir and Varzi 2021: 176)* Cotnoir and Varzi mention several examples of permissible objects, the existence of which look like reasons for some to reject permissivism:

> [C]onsider a sum of two distant stars (Eberle, 1970, p. 41), or a sum composed of a person's left foot and the carburetor of their neighbor's car (Chisholm, 1976, p. 222), or perhaps a fusion of London Bridge, certain sub-atomic particles located far beneath the surface of the moon, and Cal Ripken, Jr. (Markosian, 1998b, p. 228). The permissivist's world is replete with such things. Indeed, permissivism warrants the existence of gerrymandered wholes of all kinds: discontinuous events, such as a sum of Lennon's death and Charles' wedding (Taylor, 1985, p. 25); fusions of arbitrary qualia, such as a color, two sounds, a position, and a moment (Quine, 1951b, p. 559); and so on. Depending on one's ontology, permissivism will also licence the existence of fusions of categorially heterogeneous entities, such as a sum of you and the color blue (van Inwagen, 1987, p. 35) or

* I note, however, another kind of permissivism that Fairchild and Hawthorne (2018:46) mention "allows for a multitude of coincident material objects whenever there is any material object. This kind of permissivist standardly holds that the variety of coincident object have divergent, often radically divergent, essences."

> a sum of one of your occurrent thoughts, an attack of measles, and a lump of cheese (Geach, 1991, p. 253). (Cotnoir and Varzi 2021: 176)

But many philosophers, notably David Lewis, believe those examples, and permissivism as a view, are perfectly *compatible* with common sense and ordinary (composite) objects, and are in addition *inevitable*:

> It is compatible insofar as common sense doesn't really deny the existence of the problematic fusions; it simply 'ignores' them…And permissivism would be inevitable insofar as there appears to be no 'feasible' way of drawing a line between those fusions we find acceptable and those we don't. (Cotnoir and Varzi 2021: 178-179)

For many like Lewis who are sympathetic to permissivism, there is no conflict between objects that common sense allows and extraordinary objects that we just ignore. Whether any of these objects fall under particular *kinds* is a related but different issue. Korman notes that permissivism is neutral with respect to kind membership of objects. A composite object could fall under a familiar kind, like TABLE, but a different composite object needn't fall under some kind that is familiar. Any particle A and particle B together compose an object, according to

permissivism, but the kinds that the object falls under, if any, is a separate question.

So permissivism allows for all the ordinary objects we typically think of, like tables and chairs, but allows for many more objects as well, like sums, fusions, or composites that appear gerrymandered or arbitrary.* We can think of nihilism and conservatism as views that put restrictions on which objects exist: the former restricts objects to whatever is non-composite or atomic, and the latter is described below.

II.3.3 Conservatism

The final view considered here within object theory, *conservatism,* restricts composition to ordinary objects such as tables and trees.† For conservatism, a plurality composes an ordinary object under some condition or set of conditions that track in some way with common sense. What makes an object "ordinary" is the key challenge for conservatism.

According to Korman, conservatism is neutral with respect to at least five other views.

---

* For a thorough list of the strengths and weaknesses of permissivism, see Cotnoir and Varzi 2021, 179-181.

† Though see Korman (2019) for a conservatist social ontology of establishments.

1. "[C]onservatism is compatible with different accounts of the persistence conditions of ordinary objects and the way in which they persist." (Korman 2015: 23)

An object may or may not persist through time, and an object may or may not persist through change, but conservatism in principle does not ahead of time decide either way.

2."[C]onservatism is compatible with different views about whether and to what extent objects are mind-dependent."

Korman goes on to qualify this in the following way: "For instance, it is open to conservatives to accept the anti-realist thesis that all ordinary objects are mind-dependent: they exist only because people take them to exist." I merely note here that Korman conflates something being *mind-dependent* with something being *not real*, but throughout the project I will understand these to be separate issues.[*]

3. [W]hile conservatives are united in their commitment to rejecting sweeping eliminativist and permissivist views—like nihilism and universalism—there is still a great deal of room for variation concerning which objects exist." (Korman 2015: 24)

---

[*] For an argument against this point with respect to social objects that could extend to anti-realism more generally, see Mason (2020).

He mentions disassembled pipes and undetached parts like table-leg-complements as examples of disputed objects within conservatism.

4. "[C]onservatism is compatible with different ways of understanding the status of debates about objects."

Some conservatists may disagree over whether there are other, equally good ways of carving up the world. In other words, perhaps there are other kinds of "ordinary" objects, relative to what kinds of beings other than us might consider ordinary; perhaps what counts as a conservative, ordinary object is relative to a species.

5. "[C]onservatism is compatible with a variety of methodological outlooks." The role that intuitions play, for example, may vary among conservatists. So conservatism need not commit to particular views regarding time, mind-dependence, ontological inventory, carving nature at the joints, and methodology.

So far I have addressed ways in which many things—atoms, we suppose—may or may count as one *object*. Nihilism answers negatively, conservatism answers positively but with qualification, and permissivism gives an unqualified positive answer. In the next section I briefly take up one view that attempts to distinguish between social objects and non-social objects.

II.4. Passinsky's View of Social Objects

Passinsky (2016:1) seeks to distinguish between social objects and non-social objects in the following way:

> [T]here is a unified class of "social" objects that differ in kind from "non-social" objects, and that warrant their own metaphysical investigation. Examples of social objects include borders such as the U.S.-Canada border, units of currency such as twenty dollar bills, official documents such as my U.S. passport, sovereign states such as the United States of America, international organizations such as the United Nations, and corporations such as Google. Examples of non-social objects include natural objects such as rocks, rivers, trees and persons; and ordinary material artifacts such as tables and chairs.

So for the account above, natural objects are non-social objects, and natural objects include things like rocks, rivers, trees, and even persons.* Further, for her, social objects are indeed artifacts, but she notes that not all artifacts are social objects; on the account above, some artifacts like tables and chairs are non-social objects. In her view, all social objects are artifacts, but not all artifacts are

* Passinsky does not say whether all natural objects are social objects, but I will consider the view (whether it is Passinsky's) that all natural objects are non-social objects.

social objects. Some artifacts are made to serve some social function while other artifacts are not made for such a purpose, where "a 'social' function is a function that can only be performed or fulfilled in the context of a society (or societies)". (Ibid: 7) An artifact like a chair, for example, has a function, but "the function that chairs are made to serve is not a social function, since something can perform the function of supporting a person in a sitting position outside the context of a society". Because some chairs are made without an intent for them to perform a function within a society, there are some chairs that are not social objects, according to this view. Passinsky recognizes that this characterization of social objects depends on the tricky question of what counts as a society, and she further recognizes that it is an open question what "serving a social function" might amount to. But she claims this characterization gives at least a working knowledge of the kinds of things we should have in mind when talking about social objects.

On Passinsky's account, objects like persons, trees, and some artifacts are not social objects. Later I will question whether such objects could be considered social. Before I do so, in the next section I look at a way, in addition to composition, of unifying many things as one. We could ask whether and how we can unify many things—objects—by *kind membership* or by *the type-token relation*. For example, we could ask whether and why some objects belong to

some *social* kind. I address the topic of the kind membership relation or the functionally equivalent type-token relation generally before asking what role agents play, if any, in the kind-membership relation or the type-token relation. In the next section I take up one account that attempts to explain these relations.

II.5. Wetzel's Observation on Biological Species and Her Type-Token Model

II.5.1 Wetzel on Biological Species

In her (2009), Linda Wetzel takes up membership as it relates to *types*. For now I will assume that types and kinds are functionally equivalent in the way they unify objects. For example, I assume there is a type RAVEN with tokens of that type, and there is a kind RAVEN with members of that kind, and that tokens of the former are exactly the members of the latter.

Wetzel draws an analogy between *species* and *words*, arguing that just as there is no single feature in common for many species—where a given species is a kind or type with members—so there is no single feature in common for many words either. She further argues for the claim that "there is nothing interesting (known or unknown) that all and only members of a living species have in common other than being members of that species (i.e., no nontrivial, interesting, 'natural,' projectible property)", and she does so by running through four main approaches to answering the question of what explains membership for a species: "the *morphological*, the *genetic*, the *population*, and the *lineage*

approaches". (Wetzel 2009: 107). The *morphological* approach focuses on resemblance or similarity among species members. However, the problem with this approach as Wetzel sees it is that

> invariably there is diversity among members of the kind, that none of the characteristic properties is had by all members of the kind, and that each of the properties may be had by members of other kinds. Therefore biologists who continue to use a morphological approach—certainly the right approach in the field—employ a set of properties to characterize a given species, no one of which is necessary but some cluster of which is thought to be sufficient.

In other words, giving necessary and sufficient conditions for species membership based on resemblance or similarity looks to be more difficult than one may at first realize, and may in fact be impossible.

The *genetic* approach claims that the commonality among members of some species will be found in their common genetic coding. Intuitively, we might think that the fundamental reason a dog and a wolf fall under the same species is due to their common genetic coding. The problem here generally is that "[a] number of authors have argued that there is every reason to expect that

there is *more* diversity at the genetic level among members of a species than at the overt morphological level". Specifically,

> biologists who favor individuating species along genetic lines find that no particular gene type is necessary for an individual to be a member of a species, nor is any particular set of genes. At best species correspond to clusters of genes. So it appears that *the same situation obtains at the genetic level as obtained at the morphological level*. That all members of a species have the "same genetic code" is as much a fiction as that they all have the same morphological structure. (Wetzel 2009:108)

So according to Wetzel, we don't find some common genetic feature that all and only members of some species kind have.

Third, the *population* approach focuses on breeding (or perhaps *potential* breeding) among members of a species. This approach might give the following criterion, in the words of Wetzel: "being able to breed with every other fertile member of the opposite sex of one's species (at least, potentially) and with no other individuals, or being the offspring of two such breeders". (Wetzel 2009:109) But there are plenty of cases of interbreeding and hybrid species, both fertile and sterile, that complicate matters, and the criterion does not take into account species that reproduce asexually.

Finally, Wetzel addresses the *lineage* approach to species membership, quoting from G.G. Simpson, where "An evolutionary species is a lineage (an ancestral-descendant sequence of populations) evolving separately from others and with its own unitary evolutionary role and tendencies." (Wetzel 2009:110) But as Wetzel points out, this rules out species that are "cooked up in a lab" and have no lineage to speak of, like a species of E. coli.

So each of the four approaches is vulnerable to counterexamples. Like many others, Wetzel assumes there are biological species, but under each of the four approaches, there looks to be no single candidate feature of biological species that, for a species kind K, *all and only* members of K share. She spends the time she does on these counterexamples to set up the analogy with types of words: scientists get by just fine assuming at least some unity among biological species. Though we may not be able to come up with some unifying feature for a given species, that lack has not hindered scientific work and progress. So it seems reasonable to assume there is *something* that unifies such species. Wetzel makes the same observation for words; for any given word, it doesn't look like there is some unifying feature that all and only words of some type exhibit. Here I will simply grant Wetzel's assumption for many types, including species and words, though I question whether the argument that relies on the analogy from species to words is compelling. Before I introduce an alternative to her view, in

the next section I describe her attempt to explain what unifies tokens under their type.

II.5.2 Wetzel's Type-Token Model[*]

Wetzel presents arguably the best case for the type-token model, focusing in particular on types and tokens of *words*. I address her views on the metaphysics of words later, but for now I merely present her view on types and tokens generally. She seeks a solution to the problem that, for many types, including some natural types like species and types of social objects like words, there does not seem to be a unifying features among all the tokens of a type.

Wetzel catalogs a variety of ways in which word tokens of some type can manifest vast differences between each other while still being united under some word type. In the third chapter, she runs through various candidates for the features that all and only tokens of a *lexicographic* word type may have in common. A lexicographic word is, loosely, what she calls a word that warrants a dictionary entry.† Lexicographic words can be written, uttered, may have

---

[*] Some material here taken from Oliphint (2022).

† Thanks to Linda Wetzel via personal correspondence for alerting me to this term. See her 2009, pp. 58-71 for her helpful discussion. Although it is called a lexico*graphic* word, I assume that 1) the term comes from the inscriptive property of the dictionary entries, 2) such a word can be spoken as well, and 3) that it is a pragmatic, contingent matter that they are expressed as

various spellings and pronunciations, and may be misspelled and mispronounced. Given the wide range of objects that may token a lexicographic word, it doesn't appear as if there is some common characteristic among its tokens that could identify and predict what word type a given object is a token of. For example, if a word type like "tree" can be tokened by ink markings, sound waves, hand gestures in sign language, and so forth, it is difficult to see what common characteristic we could point to among that vast variation between tokens. She then concludes the following:

> Thus my answer to the question posed earlier, 'Is there anything all and only tokens of a particular word have in common other than being tokens of that word (i.e., any linguistically nontrivial, 'natural,' projectible property)?' is no. (Wetzel, 70)

According to this account, the only thing all and only tokens of a particular word have in common is *being tokens of that word*, and this generalizes for types and tokens beyond just words. Call the general thesis *Wetzel's Thesis*.

But an obvious objection immediately arises: on this account, whether some token is a token of a particular type appears to be merely a brute fact.

---

inscriptions in dictionaries and lexicons, and that the audiobook versions produce lexicophonic equivalences.

Wetzel's answer to this objection is that there are other factors like spelling, pronunciation, linguistic context, and so forth, "that help determine, for each word token *t*, what word type *T* it is a token of and why". But in the end her point is an epistemological one; perhaps as agents we pick up clues from tokens that help determine which type many objects token. That is not a metaphysical claim about what grounds the fact that all and only tokens under consideration are tokens of some particular type. If there is no more fundamental explanation for why some object *o* is a token of some particular word type *T*, then the fact that some *o* is a token of *T* is an ungrounded, fundamental fact, and it is difficult to see how such an ungrounded, fundamental fact is not brute; *o just is* a token of *T*, according to Wetzel's type-token account; there is no further explanation.

If there is no further explanation, then on Wetzel's account there is no role that agents play in explaining those conditions. But as I discuss later, the theory seems unsatisfying as an explanatory account of the type-token relation generally. And for objects in the social world, we intuitively think that agents play *some* role in explaining why an object is a token of a social type, or is a member of a social kind. Epstein (2015) gives a more satisfying account of kind membership that, in contrast to Wetzel's account, includes the role that social agents might play in kind membership for some kinds. In the next section I describe how Epstein includes the role that social agents play for kind

membership of *social* kinds, and I later argue that Epstein's model serves well for modeling kind membership generally.

II.6. Epstein's Two-Dimensional Model

In Epstein (2009), he puts forth a model for social kinds in reaction against two prominent models that he calls 1) *ontological individualism* and 2) the *Standard Model*. I will describe each in turn. Ontological individualism is the view that the facts about a social object are determined exhaustively by facts about the individuals or agents that make up the object. For example, according to ontological individualism, facts about the Supreme Court justices exhaustively determine facts about the Supreme Court. One of the main purposes of Epstein (2015) is to argue against that view: "Ontological individualism is false. The social facts do not supervene on the individualistic ones." (36)

Take Epstein's example of the Starbucks corporation as a social object. He examines dependence relations between facts about *the individuals involved in the corporation* and, quite separately, facts about *the corporation* as an object. One might think that the Starbucks corporation is made up only of its collective employees, but Epstein gives the following scenario as a counterexample. Suppose Starbucks was close to financial collapse and insolvency, and suppose there was a freak accident overnight at many of its stores, where a power spike caused the equipment to break and the food to go bad, which then caused the

corporation to be insolvent. None of the individuals in the corporation were involved in the incident. (Epstein 2015: 46) In that case, the facts about the corporation's transition from solvency to insolvency do not depend on any facts about individuals. Ontological individualism cannot explain this sort of example because no individual persons were involved in such a drastic change to the social object (the corporation). Yet the incident clearly affected the object in such a significant way that the object ceased to exist as a result.

Epstein, following Guala (2007), calls the other dominant model within social ontology the "Standard Model" and describes it this way:

> The idea is this: the social world is a kind of projection of our thoughts, or attitudes, onto the world. We, as a community, make the social world by thinking of it in a particular way. The bills in my pocket are money because we all think of them as money. The president has the powers he does because we grant him those powers. America is a nation because we think of it as such. The social world, quite generally, is the social world in virtue of our beliefs about it. (Epstein 2015: 50)

This model makes social objects completely or at least mostly agent-dependent, existing because of the thoughts and beliefs of social agents. The idea has precedent as far back as Hume and is found more recently in the work of Searle

and others (Epstein 2015, Ch. 4). A defining element of the Standard Model is that it includes *rules* or *conventions* set up by social agents that give conditions for objects in the social world to be what they are. On this model some community collectively accepts, for example, the border that separates their land from land owned by others, because there are complex rules or conventions that have established the social object, the border.

Epstein argues that these two models are concerned with different kinds of relations: "Ontological individualism is about one relation between individuals and the social world, and the Standard Model is about an entirely different relation". (Epstein 2015:56) Epstein refers to these relations as *grounding* and *anchoring*, respectively. Roughly, for Epstein, one fact (or set of facts) *F grounds* another *F'* if *F* serves as a "metaphysical reason" for *F'*. By contrast, anchoring is a relation between the facts about social agents and facts about the rules, conventions, or conditions in place for an object to be a member of some kind. It is "the 'putting in place' relation that holds between a set of facts and the grounding conditions for a kind". (Epstein 2015: 81)

Epstein allows for grounding and anchoring conditions to obtain in the actual world or in other possible worlds. For example, in the actual world we give conventions and rules for a baseball game, but it is possible that at least some of those rules and conventions could be different; there is a possible world

where there are four outfielders, let's say (and assuming we're still talking about the same sport). But there is, of course, more than one possible world where the conditions differ in that way. To support the notion that conventions can occur across possibilities, Epstein (2015) makes use of *frames* and *frame principles*.* A frame is a universe of possible worlds, or "a set of possible worlds in which the grounding conditions for social facts are fixed in a particular way". A *frame principle* then gives the rules for the various grounding possibilities represented by the worlds in the frame. For example, there are rules and conditions for something to be a dollar bill, and those rules and conditions can obtain in the actual world $w_@$ and in other possible worlds $w_1 \ldots w_n$ as well. Here is how Epstein puts together anchoring, frames, and frame principles:

> I will take anchoring to be a relation between a set of facts and a frame principle. For a set of facts to anchor a frame principle is for those facts to be the metaphysical reason that the frame principle is the case. In this sense, anchoring is very much like grounding. For a set of facts $g_1, \ldots, g_m$ to ground fact f is for $g_1, \ldots, g_m$ to be the metaphysical reason that f obtains in a world. For

* Frames and frame principles do not feature so prominently in Epstein's later work, for example in his (2019). I note their inclusion here because those elements are included in his (2015) explanation of the distinction between grounding and anchoring.

> a set of facts $a_1, \ldots, a_n$ to anchor a frame principle R is for $a_1, \ldots, a_n$ to be the metaphysical reason that R holds for the frame. Both are "metaphysical reason" relations. But they do different work, and stand between different sorts of relata. (Epstein 2015: 82)

In the case of a dollar bill, '[w]henever the grounding conditions are satisfied by some object, that object has the social property *being a dollar*'. (84) Facts about an object being a dollar depend on rules, conventions, collective acceptance, and so on that social agents set up, and that dependence relation is the anchoring relation.

Epstein later includes an element within his model that further explains how facts can obtain across possibilities. Take the example above of the Starbucks corporation, an *actual* object. In the example, the corporation becomes insolvent, and its insolvency is a non-actual but possible state of affairs involving that very object. Epstein's UNIVERSALITY principle makes use of anchoring, frames, and frame principles and allows for social facts to "export" to other worlds:

> (UNIVERSALITY) Given a social fact of the form ***x is K*** whose grounding conditions are anchored by actual facts $\boldsymbol{A_1 \ldots A_n}$. The fact ***x is K*** can obtain in other worlds even where $\boldsymbol{A_1 \ldots A_n}$ fails to obtain. In other words, the actual anchors set up a "universal

> tool" that can be instantiated across contexts of obtaining.
>
> (Epstein 2019: 772)*

Or take the kind HATE CRIME, for example. Suppose the grounding conditions for facts about the kind HATE CRIME are anchored by actual facts $A_1...A_n$, where those anchoring facts depend on facts about the social agents who give conditions for something to fall under the kind HATE CRIME. Now consider how that kind can obtain across times in the actual world. Suppose that facts $A_1...A_n$ did not obtain during the year 1200, i.e. that no individuals or communities during that time collectively accepted conditions for when a hate crime occurs. That does not mean from our temporal vantage point that there were no hate crimes around the year 1200. The fact that no individual or society

---

* UNIVERSALITY is similar to Thomasson's application conditions:

> Ks exist iff the application conditions for 'K' are fulfilled. But that would still leave us with the problem that there may be worlds in which Ks exist, but the term 'K' does not exist, or is associated with different application conditions. We can avoid this problem in turn by rigidifying our reference to the application conditions for 'K': it is the application conditions *actually associated with 'K'*, here and now at our world, that are relevant, regardless of the world at which the existence claim is evaluated. (Thomasson 2015: 86)

during the year 1200 was setting up conditions for the kind HATE CRIME to obtain does not negate the fact that hate crimes were in fact going on.

In the same way as the temporal example, UNIVERSALITY can apply modally across possibilities as well. Suppose further that facts $A_1...A_n$ do not obtain in w_1, because no individual or society in that world has set up conditions for the kind HATE CRIME to obtain. UNIVERSALITY permits us to take whatever conventions there are for some social kind, like the conditions for being a corporation or a dollar bill, and "peer in" to other possibilities to see whether conventions hold there.

Consider an example from (Hawley 2017) that helps further illustrate the difference between grounding and anchoring:

> The Institute of Philosophy (IP) in London offers membership both to individual philosophers and to philosophy departments of universities in the United Kingdom. Professor Gromit is a member of the Department of Philosophy at the University of Wensleydale, and the Wensleydale department is a member of the IP. Is Professor Gromit a member of the IP, in virtue of his being a member of a member of the IP? We do not try to answer this question by considering the metaphysics of social groups. Instead, we consult the website of the IP, where we find that

> members of institutional members are not automatically members of the IP; evidently, the IP could have adopted different regulations, rendering membership transitive. (Hawley 2017: 402)

What Hawley is effectively saying is that social agents, particularly the IP, give conditions or regulations for group membership, and those conditions can be found on their website. Social agents, particularly the IP, set up the anchors that determine membership, so finding out those conditions for membership merely takes a glance at the website to discover what those conditions are.

Epstein's model, particularly its inclusion of the anchoring relation, offers a clear explanation for how objects are members of social kinds, giving a more complete account than both ontological individualism and the Standard Model. Facts about agents (and perhaps other facts) anchor the rules, conventions, or membership conditions for an object to be a member of some social kind.

II.7. Conceptual Engineering and Conceptual Ethics

In their (2020) "Introduction: A Guided Tour of Conceptual Engineering and Conceptual Ethics", Herman Cappelen and David Plunkett describe their topic, *conceptual engineering* and *conceptual ethics*:

> We don't think these expressions come with fixed meanings. The previous literature has used them in different ways and so do the authors in this volume. These terms are often used without precise definitions by those working in the field. Moreover, when they are given more precise definitions by philosophers, these definitions often contradict those given by others. As editors, we could have played the terminology police for those contributing to this volume. But that would be an exercise in futility. Instead, we have decided to let a thousand (or at least a few) flowers bloom. Contributors use central terms, such as 'conceptual engineering', 'conceptual ethics', 'revision', and 'amelioration', in different ways, often explicitly so. That's how it should be given that this is currently a fast moving literature involving philosophers from many different background and sub-fields. (Cappelen and Plunkett 2020: 2)

It is important to note for my purposes here in describing the project that, as the authors mention, there are many different ideas of whatever the project is that involves conceptual engineering and conceptual ethics. Whatever that project is, I will call it "CE" for the sake of brevity. Because the authors discuss conceptual

engineering first, followed by conceptual ethics, I will briefly describe both in that order.

Cappelen captures the concern of conceptual engineering this way:

> According to Cappelen (2018), *conceptual engineering is concerned with the assessment and improvement of concepts*. However, since it's unclear and controversial what concepts are (and whether there are any), it's better to broaden the scope along the following lines:
>
> **Conceptual engineering** = (i) The assessment of representational devices, (ii) reflections on and proposal for how to improve representational devices, and (iii) efforts to implement the proposed improvements…
>
> Why call it 'conceptual' engineering when it's about representational devices more generally? Purely for aesthetic reason: 'representational devices engineering' doesn't roll off the tongue in the way 'conceptual engineering' does. (Cappelen and Plunkett 2020:3)

In the next chapter I will talk more about what representational devices ("RD"s) are, but for now think of *words* as representative examples. In fact, what kinds of

things count as RD's, whether concepts, meanings, thoughts, etc., is a point of discussion within conceptual engineering. Whatever the RD's are, conceptual engineering recognizes that some of those RD's will be better than others for certain purposes. Some RD's may have defects, and Cappelen has identified four basic strategies for improving RD's with defects: 1) do nothing, 2) abandon the RD, 3) improve the RD, or 4) replace the RD. (Cappelen and Plunkett 2020:3) How one might go about executing any of the last three strategies remains an open question for CE. Finally, in his chapter defending the project of conceptual engineering, Cappelen gives his "master argument" for it:

1. If W is a word that has a meaning M, then there are many similar meanings, M_1, M_2,. . . ,M_n, W could have.

2. We have no good reason to think that the meaning that W ended up with is the best meaning W could have: there will typically be indefinitely many alternative meanings that would be better meanings for W.

3. When we speak, think, and theorize it's important to make sure our words have as good meanings as possible.

4. As a corollary: when doing philosophy, we should try to find good meanings for core philosophical terms and

they will typically not be the meanings those words as a matter of fact have.

5. So no matter what topic a philosopher is concerned with, they should assess and ameliorate the meanings of central terms. (Capellen and Plunkett 2020: 134)

This argument will come up in the next chapter as I seek to connect the 2D model with CE.

Plunkett represents the conceptual ethics side of CE and describes conceptual ethics this way:

> Broadly, conceptual ethics concerns a range of normative and evaluative issues about thought, talk, and representation. Those include issues about which concepts we should use, ways in which concepts can be defective, what we should mean by our words, and when we should refrain from using certain words. (Which issues one thinks belong on this list, as well as how these issues are related to each other, will obviously depend on one's further philosophical commitments.) (Cappelen and Plunkett 2020: 4)

But, he says, we shouldn't think of the ethical component too narrowly:

> The use of the term 'ethics' here in 'conceptual ethics' is meant very broadly, to cover "both the study of what one should or ought to do (dually, what can permissibly be done) as well as the study of which actions and outcomes are good or bad, better or worse". Thus, this use of 'ethics' is not meant to privilege moral/political norms in particular (vs., e.g., those that find their central home in epistemology, metaphysics, aesthetics, etc.).

So the normative conceptual element of conceptual ethics is as broad as the field of ethics itself.* For example, if some community believes they should engineer some concept, then various stakeholders in the discussion will have all kinds of motivations and reasons for why the concept should be improved. Perhaps some will focus on pragmatic or consequentialist reasons for engineering the concept, while others might have different goals and duties that drive them to engineer the concept along different lines.

One of the most helpful aspects of Cappelen and Plunkett's introduction is their list of questions that are central to discussions within CE. For my purposes, I will highlight a few of these questions. For example, they ask:

---

* See also Crisp (2022), who offers an ethics *for* conceptual engineering.

> What are the objects being assessed and improved (and do they exist)? If the aim of conceptual engineering is to assess and improve concepts—or other representational devices—then we are ultimately on the hook for an account of what these objects are. (Cappelen and Plunkett 2020: 7)

If a community seeks to change a concept, a word, or some other representational device, this metaphysical question asks *what* exactly is being changed. (I will address the metaphysics of words in the third chapter.) Closely related, whatever these representational devices are, Cappelen and Plunkett ask the following: "How important is it to have a correct descriptive account of those devices in order to do the engineering project well?" (Cappelen and Plunkett 2020:14) It won't do much good if we somehow identify whatever is assessed and improved, but can't describe these things for some reason. They give the analogy of building a bridge: "[T]o think about how to improve a particular bridge, you need to know about that bridge—the ameliorative work can't be done in isolation from the descriptive work". Perhaps the analogy fails in some ways for CE, but it is obvious that the more we know about the things we are assessing and improving the better we can do the work of assessing and improving.

Even more relevant to the project here, they ask,

> to what extent can conceptual engineering change non-linguistic and non-conceptual aspects of the world? An obvious connection is this: if conceptual engineering succeeds in a particular case, it will change how people think, talk, and act on the (non-conceptual and non-linguistic) world…for those who think conceptual engineering operates on concepts and also think that some concepts (or our use of them) can be constitutive of some element of non-conceptual reality (e.g, parts of social reality), there's an interesting connection: amelioration of an important social concept can change the nature of the relevant part of social reality (since the concept is partly constitutive of some element of social reality). (Cappelen and Plunkett 2020: 12)

Later I will make the connection between the 2D model and CE more explicit, but to preview that connection we could restate what Cappelen and Plunkett say directly above in terms of the social conditions for something to be an instance of some concept. For example, there could be conditions for a person to be an instance of a social group involving race that, over time, improves in some way and therefore changes the way a community thinks about that concept.

Finally, they ask a question I will directly address in the next chapter:

> How often are we already engaged in conceptual engineering? And do we need to be aware of doing conceptual engineering in order to do it well? …[S]ome philosophers think that much of existing philosophical inquiry involves conceptual engineering to some degree. Many hold that this engagement with conceptual engineering is going on implicitly, perhaps even without the philosophers themselves being aware that is what they are up to…how much does one need to be aware of doing conceptual engineering to count as doing it? It also raises the question: how much does one need to be aware of doing it in order to do it well? Is explicit engagement with conceptual engineering always better than implicit engagement with it? (Cappelen and Plunkett 2020:14)

As I say in the next chapter, the kind of inquiry that counts under the project of CE is ongoing in the field of social ontology, so I argue there that the two fields are connected. But I also argue that what takes place under CE explicitly in some cases takes place implicitly in other works that never mention conceptual engineering or conceptual ethics by name, and I offer recent work from Ralph Wedgwood as an example.

The project of CE is as wide in its scope as the discipline of philosophy itself. The ongoing questions within CE and the disagreements over even what the project is and whether it is worthwhile reflect its wide breadth. My project here takes up only a few of the questions that Cappelen and Plunkett raise. But insofar as there is some relation between the social conditions for some concepts, including some metaphysical concepts, and the metaphysics of representational devices and social objects like words, CE as an ongoing project is worth considering, and I seek to show throughout the rest of the chapters why that is the case.

II.8. The Metaphysics of Words

If words are some of the representational devices to be assessed and possibly improved, then some of the questions Cappelen and Plunkett ask are about the metaphysics of words. There has been a growing body of literature on the metaphysics of words since David Kaplan's 1990 article "Words" sparked ongoing conversation and research involving the metaphysics of words themselves—what words are. To place in context what I say later in chapter three on the topic, I briefly summarize a few of the formative, early works from the literature on the metaphysics of words.

Kaplan asks, "How should words be individuated? What is the nature of a word?" (Kaplan 1990: 94) He also asks a more specific question: "What

determines an utterance to have the form "a=a" as opposed to the form "a=b"?" He offers a thesis of sorts:

> I am convinced that we can achieve a highly salutary clearing of the air about the nature of language, and especially about some critical differences between natural languages and logicians' idealizations, if we study the ontology and individuation of words.

Kaplan interacts with two different theories of the metaphysics of words. The conventional theory individuates words using the now-familiar type-token distinction. Kaplan calls that theory the *orthographic* conception, and he contrasts that view with his recommended theory, the *common currency* conception, described below.

Kaplan first distinguishes between *expressions* and the *occurrences* of an expression. In the examples he gives, he claims "Boston" does not contain six letters, it contains five, with two occurrences of the letter "o". In the same way with words, when yelling "Help! Help!", only one word is uttered, but there are two occurrences. For Kaplan, an utterance is not identical to the word that is used; an utterance of "Help!" is not identical to the word help, and the same principle holds for inscriptions. Utterances and inscriptions "are the physical media by which we transmit words from one to another". (Kaplan 1990: 97)

One of Kaplan's central claims is that "in interpersonal linguistic communication, words must take on a physical embodiment". So words are separate from, but can take on, physical embodiments like utterances and inscriptions. It is common then to think of words as types that can take on tokens, and that view is what Kaplan calls the *orthographic* conception, which he will argue against. Kaplan proposes the common currency conception or the *stage/continuant* model of words instead. On this view, "utterances and inscriptions are stages of words, which are the continuants made up of these interpersonal stages along with some more mysterious intrapersonal stages." (Kaplan 1990: 98) And Kaplan sees his view as naturalistic, while the orthographic conception favors a realist position regarding abstracta (because types are abstract). On the common currency conception, there can be one word with various pronunciations and spellings. And there are dialectical and idiolectical pronunciations (as in speech defects).

But Kaplan also wants to ask whether there is intrapersonal continuity with words. If someone transmits a word to me, and I transmit that word to someone else, have I retained the identity of the original word such that I am transmitting the same word to someone else? Kaplan believes identity retention for a word is possible through transmission, and offers a condition where, in transmission, "the difference in sound or shape or spelling, can be just about as

great as you would like it to be". (Kaplan 1990: 101) For Kaplan, part of answering the question of word identity through transmission involves answering what happens in the "black box" (the mind) when receiving input and transmitting output. And Kaplan believes we cannot know what goes on in the black box, but we can observe behavior that may give us clues. If I tell someone to repeat a name and I observe that they repeat the name successfully every time, I can describe the person's output as a repetition of the name.

For Kaplan, something is clearly going on in the black box, but he withholds any commitment to what exactly goes on psychologically and what kind of thing the mind is. But there is input into the ear, some sort of pattern recognition in the black box, and output through the mouth that matches—in some way and in varying degrees—the input. And some may have barriers to their imitative ability, like a hearing problem or speech defect, where the input or output does not resemble what is received or transmitted. But Kaplan wants to say that even in the cases where the physical properties do not resemble between input and output, the subject is repeating the word. So he says:

> The identification of a word uttered or inscribed with one heard or read is not a matter of resemblance between the two physical embodiments...Rather it is a matter of intrapersonal continuity, a matter of intention. (Kaplan 1990: 104)

It is intention that "decides the matter" of word identity.

Kaplan is pre-theoretically committed to a naturalistic view of words, which rules out abstract types. What differs on his account from the causal chain theory is the emphasis and location of names, for example, inside the mind. One view of names is that they "are not part of any language, because one can just create new names at will, whereas the lexicon of words other than names forms a relatively stable body characteristic of the language". (Kaplan 1990: 113) On this view, common currency names can be created at will, but generic names cannot.

Kaplan's 1990 paper spawned a research project into the metaphysics of words, but it has only recently generated a consistent stream of literature in the past five years or so. Cappelen (1999) took up Kaplan's topic and offered a few challenges to it. Like Kaplan, Cappelen observes that the conventional view uses the type-token distinction to individuate or metaphysically describe and categorize words:

> (1) There is a distinction between types and tokens. Types are abstract objects, for instance patterns. Tokens are concrete particulars, for instance sounds or inscriptions. A particular is a token of a type, T, if and only if it stands in some appropriate

> relation to T. If T is taken to be a pattern, the relevant relation is the instantiation relation.
>
> (2) The abstract entities that are sign types are semantically and syntactically neutral objects. They have their semantic and syntactic properties contingently, through conventions. (Cappelen 1999: 92)

This Type-Token model is rejected by Kaplan in favor of his common currency conception, renamed by Cappelen as an *intentional* theory of words. For Cappelen,

> A theory of words is intentional if it says that it is a necessary or sufficient condition for something being a token a word that the producer of the token was in a certain intentional state at the production time (intentionalist theories differ in how they describe the required intentional state). I call the claim that this is a necessary condition the Necessity Thesis and that claim that it's a sufficient condition the Sufficiency Thesis.

In the first part of the paper, Cappelen describes Kaplan's view and offers counterexamples to the intentional theory. I won't repeat Kaplan's view here, except to note Cappelen's relevant commentary. Cappelen describes Kaplan's

word-object—the thing that is created by a person and continues through interpersonal and intrapersonal stages—as a "very strange thing".

If intentions are sufficient for something to be a particular word, then an object that has no resemblance to the standard word can count as a particular instance of a word. But that lack of resemblance is, for Cappelen, an unacceptable and counterintuitive consequence of an intentional view. Now, although we do at times accept tokens that do not resemble conventional signs, and do accept them for reasons involving the producer's intent, Cappelen believes this does not give us license to accept just any token as a repetition of the original. So the Sufficiency Thesis fails.

Cappelen also offers a counterexample to the Necessity Thesis. Say I found a piece of paper with

CAN YOU SPARE A QUARTER?

and use it to ask people for a quarter (while remaining silent). If the token was produced with wrong intentions or no intentions, then on the Necessity Thesis the token isn't a sentence and it was never used to ask what we otherwise would think it's asking. Those consequences are unacceptable for Cappelen. He wants to distinguish intentional*ism*, which identifies words with intentions, from the

"view that some intentional (linguistic) activity is required for there to be sign tokens". (Cappelen 1999: 99) It is this latter view that he wants to defend against both intentionalism and the type-token model. First, he observes that the words on the page are tokens of words because of convention, which is contingent; had history gone another way, other tokens would be conventional words. And convention involves intention, in that conventions "are upheld by intentional linguistic activity".

For Cappelen, there are two kinds of conventions:

> (C1) Entities with such-and-such properties count as tokens of the same sign
>
> (C2) Entities with properties P (where entities with P count as tokens of the same sign according to some Cl-type convention) count as tokens of the same sign as tokens with properties P' (where P' entities count as tokens of the same sign according to some Cl-type convention)

C1-type conventions and C2-type conventions are related in the following way:

> Cl-type conventions determine the same-sign relation within a sign system (such as spoken, written, Braille, Semaphore etc),

> while C2-type conventions determine when tokens in different sign systems are tokens of the same sign.

The disagreement between intentionalists and non-intentionalists turn on what kind of properties the "such-and-such" properties are in C1: intentionalists locate the properties in intentions, while non-intentionalists deny such a claim.

For Cappelen, the type-token model for words does not take conventions into account, but should. The type-token model "treats expression types as paradigmatic abstract objects", e.g. patterns of some kind. And he explains the problem with the model this way:

> if the same-sign relation is conventional, then either sign-types are abstract objects that can only exist through conventions (which would disqualify for instance patterns) or the instantiation relation is conventional. The conventional element of the same-sign relation must be introduced, but is missing from the traditional model. (Cappelen 1999: 100)

So Cappelen offers challenges to both the type-token model and Kaplan's intention-based alternative.

Alward (2005) picks up on the exchange between Kaplan and Cappelen, and opens with a comment on the literature from Kaplan's 1990 piece up to 2005: "The critical reaction to Kaplan's "Words" in the decade or so since its

publication is better described as a trickle than a torrent". (Alward 2005: 172) If, as Cappelen believes, the ontology of words has seen little attention before Kaplan's piece, and if, as Alward believes, there was little reaction to Kaplan's paper, the literature on the ontology of words up to 2005 remains scarce.

Much of Alward's paper merely describes and offers comments on Kaplan and Cappelen, so I won't repeat that material. A few years later, Epstein (2009) enters the discussion on the metaphysics of words that will later develop more broadly into his model for social ontology. Much of the subsequent literature takes the type-token model and Kaplan's response as a base for discussion.[*] But more recently there has been some work that takes on a more constructive character, independent from the initial discussions involving Kaplan and his reaction to the type-token model. Gasparri (2020), for example, argues that there may be many acceptable, non-competing theories of the nature of words and how to individuate them. Carrara and Smid (2022) seek to combine classical mereology with plural logic and apply those fields to word types, because word types have been understood by some to be a counterexample to extensional mereology. And my own work in Oliphint (2022)

[*] See Hawthorne and Lepore (2011), Bromberger (2011), O'Madagain (2014), Balletta (2019), and Miller (2019).

attempts to show the superiority of Epstein's model for social ontology over the type-token model for explaining the puzzling metaphysical features of words.

II.9. Conclusion

To show the utility of the two-dimensional model for unifying objects into kinds, I first distinguished syntactic unity and semantic unity, drawing on Cotnoir (2013). Because there are several views about what counts as an object, I then described the major theories that seek to give an account of what counts as a composite object. I leave the question of which theory is most consistent with social objects until the next chapter, but I introduce Passinsky's (2016) account of what makes an object social, i.e. what objects are members of a social kind. Through her type-token model, Wetzel (2009) attempts to give an account for a particular social object, a word, and I questioned whether the model provides a satisfactory answer. I sketch Epstein's (2009) model for kind membership as what I will argue is a more satisfactory model for unifying objects. I then describe the project of conceptual engineering and conceptual ethics (CE). The project makes use of representational devices like words, so I then highlighted a few influential pieces in the literature on the metaphysics of words. In the next chapter I connect the project of CE and the 2D model and then apply it to the main object theories, assessing which concept of OBJECT is most friendly to an ontology of the social world.

II.10. References

Alward, Peter. (2005). "Between the Lines of Age: Reflections on the Metaphysics of Words." *Pacific Philosophical Quarterly* 86 (2): 172–87. https://doi.org/10.1111/j.1468-0114.2005.00221.x.

Balletta, Sandro. (2019). "The Metaphysics of Words." *Theoria* 85 (1): 31–48. https://doi.org/10.1111/theo.12174.

Berto, Francesco, and Massimiliano Carrara. (2009). "To Exist and to Count: A Note on the Minimalist View." *Dialectica* 63 (3): 343–56. https://doi.org/10.1111/j.1746-8361.2009.01198.x.

Bromberger, Sylvain. (2011). "What Are Words? Comments on Kaplan, on Hawthorne and Lepore, and on the Issue." *Journal of Philosophy* 108 (9):486-503.

Cappelen, Herman. (1999). "Intentions in Words." *Noûs* 33 (1): 92–102. https://doi.org/10.1111/0029-4624.00143.

Cappelen, Herman and Plunkett, David. (2020). "Introduction: A Guided Tour of Conceptual Engineering and Conceptual Ethics" in *Conceptual Engineering and Conceptual Ethics*. Alexis Burgess, Herman Cappelen, and David Plunkett, eds. Oxford University Press.

Carrara, Massimiliano, and Jeroen Smid. (2022). "There Is Overlap in Letters: A Mereological Definition of Word Types." *Synthese* 200 (2): 178. https://doi.org/10.1007/s11229-022-03664-4.

Cotnoir, Aaron J. (2013). "Composition as General Identity" in *Oxford Studies in Metaphysics*. (eds.) K. Bennett and D. Zimmerman, vol 8. Oxford: Oxford University Press, 295-322.

Cotnoir, A.J. and Achille C. Varzi. (2021). *Mereology*. Oxford: Oxford University Press.

Crisp, Roger. (2022). "Towards an Ethics of Conceptual Engineering." *Inquiry*, 1–14. https://doi.org/10.1080/0020174X.2022.2075454.

Epstein, Brian. (2015). *The Ant Trap: Rebuilding the Foundations of the Social Sciences*. Oxford University Press.

Epstein, Brian. (2019). Anchoring versus Grounding: Reply to Schaffer. *Philosophy and Phenomenological Research* 99 (3): 768–81. https://doi.org/10.1111/phpr.12644.

Gasparri, Luca. (2019). "A Pluralistic Theory of Wordhood." *Mind & Language*. https://doi.org/10.1111/mila.12297.

Guala, Francesco. (2007). "The Philosophy of Social Science: Metaphysical and Empirical." *Philosophy Compass* 2 (6):954–80.

Hawley, Katherine. (2017). "Social Mereology." *Journal of the American Philosophical Association* 3 (4): 395–411. https://doi.org/10.1017/apa.2017.33.

Hawthorne, John & Lepore, Ernest (2011). "On Words." *Journal of Philosophy* 108 (9):447-485.

Horden, John, and Dan López de Sa. (2020). "Groups as Pluralities." *Synthese*, August. https://doi.org/10.1007/s11229-020-02715-y.

Kaplan, David. (1990). "Words." *Aristotelian Society Supplementary* Volume 64 (1): 93–120. https://doi.org/10.1093/aristoteliansupp/64.1.93.

Korman, Daniel Z. (2015). *Objects: Nothing Out of the Ordinary*. Oxford: Oxford University Press.

Korman, Daniel Z. (2019). "The Metaphysics of Establishments." *Australasian Journal of Philosophy* 98 (3): 434–48.

Mason, Rebecca. (2020). "Against Social Kind Anti-Realism." *Metaphysics* 3 (1): 55–67. https://doi.org/10.5334/met.30.

Merricks, Trenton. (2001). *Objects and Persons*. Oxford: Oxford University Press.

Miller, J. T. M. 2019. "A Bundle Theory of Words." *Synthese*, October. https://doi.org/10.1007/s11229-019-02430-3.

Oliphint, Jared S. (2022). "Using a Two-Dimensional Model from Social Ontology to Explain the Puzzling Metaphysical Features of Words." *Synthese* 200 (3): 227. https://doi.org/10.1007/s11229-022-03716-9.

O'Madagain, Cathal. (2014). Indexicals and the Metaphysics of Semantic Tokens: When Shapes and Sounds become Utterances. *Thought: A Journal of Philosophy* 3 (1):71-79.

Passinsky, Asya. (2016). *The Metaphysics of Social Objects*. Dissertation. New York University.

Passinsky, Asya. (2020). "Social Objects, Response-Dependence, and Realism." *Journal of the American Philosophical Association* 6 (4): 431–43. https://doi.org/10.1017/apa.2019.51.

Thomasson, Amie L. (2015). *Ontology Made Easy*. Oxford: Oxford University Press.

van Inwagen, Peter. (1990). *Material Beings*. Ithaca: Cornell University Press.

Wetzel, Linda. (2009). *Types and Tokens: On Abstract Objects*. MIT Press.

CHAPTER III

THE TWO-DIMENSIONAL MODEL OF KINDS WITH CONCEPTUAL ENGINEERING AND CONCEPTUAL ETHICS

In the previous chapter I outlined two models that attempt to explain why, in some cases, many things are instances of a type or kind: Wetzel's type-token model and Epstein's two-dimensional ("2D") model. I also described the projects of conceptual engineering and conceptual ethics (together "CE"), and highlighted questions from those projects that have to do with metaphysics broadly and social ontology. Below I argue that 1) the 2D model is superior to the type-token model for explaining how many things can be so unified, and 2) that the model's superiority is increased further when the model incorporates some of the features from the project of CE. I then apply the model to object theories to clarify what I will mean by "object", arguing that permissivism is most consistent with an ontology of the social world. But first I show the superiority of the 2D model over the type-token model.

III.1. The Two-Dimensional Model of Kinds Over the Type-Token Model

Recall from the previous chapter Wetzel's theory that, for many types, what feature unites tokens under a single type is *being tokens of that type*. Wetzel

believes her theory follows from the fact that, for many types of objects (e.g. biological species and words, for example), there is no single, unifying feature that all and only tokens of that type exhibit. As I mentioned in the previous chapter, what follows from her theory is less than satisfying: that some token is a token of some type is an ungrounded, fundamental fact. It is difficult to see how such an ungrounded, fundamental fact is not brute.

I do not deny the observation that motivates Wetzel's theory: for many types, there does not seem to be a single criterion or feature that can unite all and only tokens of that type. But it would be preferable if we could avoid following Wetzel's conclusion that, for many types, the feature that all and only tokens of some type have is merely being tokens of that type. Having no further explanation for why some object falls under a particular type or kind, beyond its being an ungrounded, fundamental fact, is less than satisfying. It would be better to find a more satisfying explanation for why an object is a token of some type. On the other hand, if we cannot find what unites those objects in the features of the *objects* themselves, where can we look?

I argue that in many cases we look in part to facts about *social agents* to explain the fact that many objects are unified under a type, kind, or concept. To account for the role that social agents play in whatever unites many objects, we need a model that takes into account the difference in relations between facts

about *the objects themselves* and facts about *the social agents who give conditions* for the unification of many objects. As I mentioned in the previous chapter, Epstein (2015; 2019) has provided such a model, and Brouwer (forthcoming) has helpfully called Epstein's model a *two-dimensional* model.* Recall that one dimension of the model involves *grounding* of social facts (where a social fact is about a social object of some kind, however defined):

> The facts that can play the grounding role with respect to social facts are not of any specific type: what they look like depends on the kind of fact that's being grounded. That something is a dollar bill is grounded in how and where it was manufactured (a fact about its history); that someone is popular is grounded in how other people regard them (a fact about attitudes); that Jimmy can't go on the roller-coaster is grounded in his height (a fact about physical properties).

But when we ask for an explanation of why *those* facts are the grounding facts, we are asking about a different explanatory relation. For example, we could ask what explains the fact that the dollar bill being manufactured by the Bureau of Printing and Engraving grounds the further fact that some rectangular piece of

* Some material in this section is taken from Oliphint (2022).

cloth is a dollar bill. That question is asking about something different than the grounding relation. Language about the grounding relation can refer to the grounding *relation*, the grounding *role* (like in the quote above), we can use *ground* as a verb, or *grounding* as a verb. Language about the anchoring relation runs the same way: we can refer to the *anchoring* relation, which is the relation between the fact that some objects are members of a kind and the facts about what Epstein calls the *anchors*, which are "the customary practices, the enactment of statutes, the judgments in historical cases, etc., that socially construct [or "anchor"] the kind" in question. Brouwer describes the difference between grounds and anchors this way :

> Now we're asking why the grounding relations run as they do. In Epstein's theory the grounding principles that govern a given case are explained and determined by different facts which he calls anchors…Anchors are often social facts themselves, and what they anchor are grounding principles of the general form 'if A, then A grounds B' (where B is some social fact). The facts that show up among the grounds and the anchors are not metaphysically different kinds of facts, and one and the same fact could feature in either role on occasion. *But the explanatory roles are different*. (Brouwer forthcoming; my italics)

The fact that some object is an instance of a kind depend on both grounds and anchors, each with a different explanatory role. An example of an anchoring explanatory role is the following: the fact that an object's causal history is a condition for something to be a dollar bill is anchored by facts about social agents who have set up such rules and conditions for dollar bills. The anchors are the social facts, including facts about social agents, that set up those conditions. Brouwer gives a few helpful examples of anchors:

- Facts concerning speech acts or inscriptions. The fact that a crime of robbery needs to involve a threat of violence is explained (in part) by the fact that a certain form of words defining the act of robbery appears in the law code.

- Facts about the functional roles of social kinds or institutions. The fact that waving your arm can constitute a greeting but thinking happy thoughts cannot is explained (in part) by the fact that the latter is not outwardly perceptible and hence not apt to play the role of a greeting.

- Facts about paradigm instances of a social kind. That dancing the Macarena involves such-and-such moves is explained by the appearance of those very moves in the 1996 music video.

- Stable patterns in behavior. The fact that deer musk-markings constitute territorial claims is explained by the fact that deer are stably disposed to react to them in a certain way. (Brouwer, forthcoming)

From this 2D model we can more clearly see the problem with a one-dimensional model like the type-token model. As we saw above, in the type-token model, when an object is said to be a token of some type, there is no fundamental explanation for *why* it is a token of *that* type. There is no fundamental explanation because a one-dimensional type-token model lacks an anchoring relation that can explain *agent-dependent conditions* for being a token of some type, or for being a member of some kind.*

In the absence of this second dimension, we are stuck with the brute fact that some token simply belongs to some type, because there is no common feature that all and only tokens of that type share. Wetzel considers the example of a grizzly bear:

> Not all adult grizzlies are big, not all are brown, not all have humps, and so forth. Almost any generalization about all grizzlies will be false if there is one. (Wetzel 2009: 70)

* Types, kinds, and concepts will again be functionally equivalent *universals* here.

If Wetzel is correct, there is no characteristic that will be true about all and only grizzly bears. If this is all there is to say about how an object can be an instance of GRIZZLY BEAR, we are left without a ground for characterizing some object in question as an instance of GRIZZLY BEAR.

But there is more to say about examples like these: the 2D model offers an explanation, for example, why members of GRIZZLY BEAR fall under that kind, yet each of those members lack some single characteristic we can point to that unites all and only grizzly bears. In addition to facts about the objects themselves (like facts about the grizzly bears in question), the 2D model allows for whatever *social* facts anchor the conditions for something to be a grizzly bear. Those social facts will be evaluated by biologists and others, and for many kinds it may not even be clear what relevant facts are given by social agents that anchor the conditions for kind membership. Part of the significance of the 2D model is that it allows for such social facts as parts of a full explanation for how several things count as instances of a kind. Another part of the model's significance is how it enhances the ongoing project of CE, discussed in the next section.

III.2. Common Ground Between CE and the Two-Dimensional Model

The superiority of the 2D model becomes even more clear when linked to the project of conceptual engineering and conceptual ethics. I will explain why

that's the case, but because reference to "concepts" appears throughout this section and the chapters that follow, I need to first specify the kinds of things I have in mind when speaking of concepts. Briefly, the concepts I have in mind are abstract universals that have instances. I will at times follow the language used by those working in the field of CE, like Cappelen and Plunkett, who talk about engineering concepts. Strictly speaking I think about the practice of engineering concepts as *selecting* some concept from among the abstract plenitude that suits whatever aims and purposes are in question. But to be consistent with the language used in the CE literature, I will use the conventional language that speaks about concepts as things that can be engineered.

Recall from the first chapter Cappelen's (2020) master argument for the project of conceptual engineering:

1. If W is a word that has a meaning M, then there are many similar meanings, M_1, M_2,. . . ,M_n, that W could have.
2. We have no good reason to think that the meaning that W ended up with is the best meaning W could have: there will typically be indefinitely many alternative meanings that would be better meanings for W.

3. When we speak, think, and theorize, it's important to make sure our words have as good meanings as possible.
4. As a corollary: when doing philosophy, we should try to find good meanings for core philosophical terms and they will typically not be the meanings those words as a matter of fact have.
5. So no matter what topic a philosopher is concerned with, they should assess and ameliorate the meanings of central terms. (Capellen and Plunkett 2020: 134)

Now consider Cappelen's argument above in the context of the following illustration from Epstein (2019). Epstein uses the concept WAR CRIMINAL to highlight the important differences between the grounding relation and the anchoring relation.* The example is worth quoting in full:

> Suppose you are a judge at the International Criminal Tribunal for the former Yugoslavia (ICTY), and Ratko Mladic is brought before you, accused of war crimes. The prosecution argues that Mladic is a war criminal, in virtue of (among other atrocities)

* Epstein actually considers the *kind* WAR CRIMINAL, but recall my assumption above that kinds and concepts play a similar enough functional role to one another to be substituted in the way I do here.

> having ordered the massacre in Srebrenica in 1995. Consider some claims Mladic's lawyers might make in his defense:
>
> G1. There was no 1995 Srebrenica massacre.
>
> G2. Even if the massacre did take place, Mladic did not order it.
>
> G3. Even if Mladic did order the massacre, he was a soldier following the orders of higher-ups.
>
> G4. Mladic's actions do not fall under the jurisdiction of the ICTY…
>
> These are claims the defense might argue to persuade the court that Mladic is not a war criminal. If one or more of these claims were true, the argument would be, Mladic does not meet the conditions for being a war criminal.

The purported facts G1-G4 are meant to ground the fact that Mladic does not instantiate the concept WAR CRIMINAL. If these purported facts hold up, so the illustration goes, Mladic does not satisfy the conditions for falling under that concept. But there is more to say about the conditions for WAR CRIMINAL:

> There is also a different strategy the defense might pursue. In addition to arguing that Mladic fails to meet certain conditions,

> the defense could further argue over *what it takes* to be a war criminal. His lawyers might argue:
>
> WC1. Perpetrating such a massacre is not a war crime.
>
> WC2. Ordering such a massacre to be perpetrated is not a war crime.
>
> WC3. If a soldier performs an action following orders from higher-ups, that action is not a war crime.
>
> WC4. Only actions that fall under the jurisdiction of a war crimes tribunal can be war crimes…
>
> This second set of claims are about what it takes—that is, what the conditions are—to be a war criminal. They are not more general or "structural" than the previous set, but are claims about the boundaries of *war criminal*, as opposed to claims about whether Mladic's actions fall within those boundaries. (Epstein 2019: 768-769)

The first set of claims involve the question of whether some object (Mladic) falls under some concept (WAR CRIMINAL), due to some facts about the object under consideration. The second set of claims is quite different, *assessing what the anchoring conditions are or should be* for some object to fall under a concept. The

assessment of those conditions involving the anchoring relation are, I claim, central to the very tasks of conceptual engineering and conceptual ethics. What takes place in this example in the case of the second strategy is the kind of task Cappelen has in mind when proposing his master argument, if we substitute *meaning* with *concept*.

Recall from the previous chapter that representational devices could include words, concepts, meanings, and so forth. The concept WAR CRIMINAL, or the meaning of "war criminal", could include any combination of WC1-WC4, depending on how we engineer the concept, and we could present arguments for what the concept WAR CRIMINAL *should* be according to some set of normative, ethical principles for the concept.[*] Let this serve as just one example on the way to answering some of Cappelen and Plunkett's questions we saw in the first chapter: "How often are we already engaged in conceptual engineering? And do we need to be aware of doing conceptual engineering in order to do it well?"

We can generalize from Epstein's example of WAR CRIMINAL: there is some universal like a type, or a concept, or a kind, and there are the conditions

[*] We could also distinguish what we might call *concrete* engineering, where we change something in the world and thereby change whether some object instantiates a given concept. Such engineering might involve the *implementation* of concepts that Cappelen and Plunkett describe (see Appendix). I owe this idea of concrete engineering to Robert Garcia.

for many things to be unified under that single universal type, concept, or kind. (In the present context where I discuss CE, I will simplify things from here on by referring only to concepts rather than types or kinds.) For many concepts, agents will play some role in engineering what the conditions are for things to be instances of a concept in question, where the concept is represented by some device like a word or phrase. And in many cases there will be better and worse conditions for some concept to have instances; if there is a concept that is defective in some way and needs improvement, the project of conceptual ethics is positioned to help highlight the normative and evaluative issues at play for concepts.

Consider the example of GRIZZLY BEAR that Wetzel mentioned above, where we observed the superiority of Epstein's 2D model to the type-token model because of the way it can explain the role of social agents for what counts as an instance of a kind. For the functionally equivalent concept, note again that biologists will come up with many (anchoring) conditions for an object to count as an instance of the concept GRIZZLY BEAR, perhaps having to do with some combination involving DNA, origin and lineage, and so forth. And there may be better and worse conditions that biologists might give for something to count as an instance of the concept GRIZZLY BEAR. Generalizing from this particular example, the projects of conceptual engineering and conceptual ethics focus on

what the anchoring conditions are for something to be an instance of a concept in question like GRIZZLY BEAR.

Note the philosophical work that the 2D-CE model is not intended to do: it does not yield, for a concept in question, whatever objective feature(s) an object must have to count as an instance of the concept. If Wetzel is correct, for many concepts like GRIZZLY BEAR there will not be a feature that all and only instances of the concept have. What the model is intended to do is to distinguish among the different relations (grounding and anchoring) between facts that are involved in the conditions given by social agents for something to be an instance of a concept, to identify the task of conceptual engineering social agents are engaged in when considering various anchoring conditions for a concept, and to identify the task of conceptual ethics when there are normative factors at play relative to a set of goals that govern what the instances of a concept should be. In fact, problems have occurred within the field of biology due to the fact that in many cases there is no such feature that all and only members of a species have:

> [F]or a discipline aiming to impose order on the natural world, taxonomy (the classification of complex organisms) is remarkably anarchic. There is reasonable agreement among taxonomists that a species should represent a distinct evolutionary lineage. But there is none about how a lineage

> should be defined. 'Species' are often created or dismissed arbitrarily, according to the individual taxonomist's adherence to one of at least 30 definitions. Crucially, there is no global oversight of taxonomic decisions — researchers can 'split or lump' species with no consideration of the consequences. (Garnett and Christidis 2017: 25)

If the quote above is correct, social agents who are biologists are in part in the business of defending their preferred set of grounding conditions for species membership. For some species, what explains some set of grounding conditions for a given species is not the features of objects alone, but also the decisions of biologists as they engineer a species concept based on normative goals. Consider the following specific anchoring conditions that are debated among biologists (and note the use of "concept"):

> The lack of universal rules for taxonomy has many consequences. A major issue is differences in tradition between classes. Many mammalian taxonomists use the phylogenetic species concept (PSC): two populations are listed as distinct species if they have a common ancestor but differ physically or genetically. Meanwhile, many bird taxonomists favour the more conservative biological species concept — the idea that true

> species should not normally produce fertile hybrids. An estimate published last year suggests that the number of bird species would more than double were bird taxonomists to adopt the PSC. (Ibid: 26)

The article goes on to mention that these classification issues matter because of normative goals involving conservation: "Depending on which species concept is used, one class can seem more threatened than another, and so receive a bigger slice of conservation funding". For many species, an explanation for why an object counts as an instance of a species concept will involve grounding conditions, social agents who engineer those grounding conditions as opposed to others, and a set of ethical, normative goals for a species concept. A model that includes types and tokens alone cannot offer such explanatory complexity.

Consider another example that illustrates the conditions that we as social agents put in place for a concept. In his chapter "Against 'Evidence'", Ralph Wedgwood (manuscript) argues that "'evidence' is a dangerous word to use in epistemology". I am not interested here in the truth or falsity of that particular claim. Instead, I am interested in what Wedgwood is doing when he examines the *concept* EVIDENCE and its representational devices such as the word 'evidence'. He states his purpose in the beginning of the chapter as follows:

> I shall investigate the concepts that the word 'evidence' can express, on the assumption that it stands for something with the kind of normative significance that epistemologists are interested in. We shall find that although we can stipulate technical senses for the term 'evidence' which will allow it to play this sort of normative role, it is doubtful whether the term has any such sense in everyday language.

Note that neither the term "conceptual engineering" nor the term "conceptual ethics" appears in the chapter. But we can assess whether Wedgwood is (perhaps unintentionally) engaging in the activity of conceptual engineering and arguing for a normative point involving conceptual ethics. I claim that that is what is going on in Wedgwood's chapter on evidence, for the following reasons.

Among the concepts that the representational device "evidence" can express, Wedgwood distinguishes between *technical* senses or concepts on the one hand, and senses or concepts of the word in *everyday language* on the other. If conceptual engineering assesses the best possible concepts for representational devices, and in turn highlights which concepts are defective or less than ideal in some way, then Wedgwood is doing just that with the concept of EVIDENCE. Perhaps the concept involves what is 'given' epistemologically for someone who believes some proposition *p*. Or perhaps the concept involves 'what is evident'

to a believer. Or perhaps on a Bayesian conception a proposition *p* is evident if it has a probability close to 1. These are concepts that Wedgwood considers, and he finds that all these concepts, as well as whatever the everyday English concept is, run into problems. So for Wedgwood there are several related concepts that are problematic, all of which are represented by the word 'evidence'. It may be that there is no concept of EVIDENCE that avoids such problems.

The fact that the concept of EVIDENCE is defective in at least some ways carries with it normative implications, according to Wedgwood. I mentioned above that he believes the term is dangerous for use in epistemology, because the use of the term "continually makes highly questionable assumptions seem much more plausible than they should… Epistemologists would do well to be much more wary in their use of this term than they have typically been". This claim seems like an instance of some of the primary concerns within conceptual ethics if anything is, given Plunkett's characterization of conceptual ethics:

> The use of the term 'ethics' here in 'conceptual ethics' is meant very broadly, to cover "both the study of what one should or ought to do (dually, what can permissibly be done) as well as the study of which actions and outcomes are good or bad, better or worse". (Cappelen and Plunkett 2020: 4)

In the quoted section above from Wedgwood, he concludes with a normative recommendation that epistemologists *should* ("would do well to") avoid talk of evidence because of all the problems he has uncovered. Such a recommendation is at least possible to follow, though the logistics of phasing out such a common word altogether may be less than probable. We could consider other examples. For these concepts and many others, the combination of the 2D model with the project of CE—what I will call the "2D-CE" model—is a tool that allows us to take into account the anchoring role that social agents play in *engineering* the grounding conditions for a concept to be instantiated, and the normative/ethical considerations for engineering some concept in question.

So far I have been speaking about what objects may count as instances of a concept. But we saw in the first chapter that there are several different theories for what counts as an object, so I need to clarify what I mean when speaking about an *object* counting as an instance of a concept. I seek here 1) to show how the 2D model and CE (together the "2D-CE" model) highlight the conditions under various theories for something to count as an object,* and 2) to examine which object theory is most compatible with an ontology of the social world. Many objects in the social world are highly valued financially and morally: an

* I have in mind here *material* objects and will use the term "object" as shorthand for "material object", but what I say below might also apply to abstract entities.

instance of money, like a dollar bill, is a social object with financial value, and political groups of people like races and genders have moral value. Because of the high value we place on objects like these, an object theory that is most compatible with objects in the social world would in principle allow us to keep such valuable objects in our ontological inventory. That would be a benefit for such an object theory. I argue that, among the object theories of nihilism, conservatism, and permissivism, permissivism is most compatible with an ontological commitment to objects in the social world.*

Now that we have covered the 2D model, the project of CE, and how both are similar, I will briefly summarize the combined 2D-CE model before applying it to the object theories covered in the first chapter. Recall Wetzel's observation that, for many types of objects, including natural types like biological species and social types like words, there is no single feature that all and only tokens of those types have in common.† For many types, there doesn't seem to be a feature of *the objects* (or "tokens") that all and only the tokens under some type have in common. For tokens of natural types like biological species and artifacts like

---

* It may actually turn out that a different object theory that is incompatible with social objects is eventually shown to be the indisputably true theory. In that case (and assuming such a theory's incompatibility with permissivism), permissivism could no longer claim as a benefit the compatibility with social objects (in the scenario we are considering). But an object theory that is incompatible with social objects in that way doesn't look forthcoming.

† Note that the condition of having at least one feature in common with all and only objects of a type is in many cases merely an *assumed* condition for unity. I neither endorse nor challenge that assumption here.

words, for example, if we ask what grounds the fact that some token is a token of a particular type, we are left without an answer. If there is anything that unites all and only such objects, we need to look somewhere other than the features of the objects.

In its most general form, the 2D-CE model includes whatever universal suits the aims and purposes of giving conditions for some object or objects to be an instance of the universal in question. I mentioned above that I take types, kinds, and concepts to be functionally equivalent. For better or worse, the project of CE generally speaks of *concepts* as the universals involved in engineering (i.e. assessing the instantiation conditions). But types, kinds, meanings (in the sense that Cappelen speaks of above in his master argument), or properties may in many cases serve the purposes of the universal in question whose instantiation conditions can be assessed. In the case of Payton (2022), for example, she speaks about social properties like the property of *being cool, being a woman,* or *being charismatic*. But she could have written the paper using concepts like COOL, WOMAN, or CHARISMATIC. Consider the central role language plays in her argument for what counts as a mark of the social. If she is right, "the literature on social ontology has devoted most of its attention" to the relationship between predicates and properties for cases like the ones above (e.g. *being cool*). One could make a similar observation for the literature on conceptual

engineering as well; concepts from the social world like WOMAN feature throughout the literature.

Below I apply the 2D-CE model to object theories with two goals in mind. First, I use the 2D-CE model to clarify what I mean by "object" when I talk in the next chapter about objects having social features. Second, I seek to assess which object theory is most compatible with social ontology. I mentioned above that many social objects are highly valued, so it would be a virtue of a theory if it allowed such valuable objects in its ontological inventory. After considering nihilism and conservatism, I argue that permissivism is most compatible with an ontology of the social world.

III.3. The 2D-CE Model Applied to Object Theories and A Case for Permissivism

III.3.1 Conditions for OBJECT Under Nihilism

Recall that nihilism is a species of eliminativism, where eliminativism restricts when composition occurs based on some specified set of conditions. For nihilists (assuming atomism), the only objects that exist are *atoms* (sometimes called *simples* or *particles*), which have no proper parts. Under this view, there are no composite objects at all: atoms A and B never compose a further object C.

According to nihilism, the only instances of the concept OBJECT are atoms: the fact that A is an object is grounded by the fact that A has no proper

parts. The fact that a purported object C is not an object is grounded by the fact that C, were it to exist, would have proper parts. But just like we can ask what the best conditions are for something to count as an instance of WAR CRIMINAL, we could ask whether *that condition*—the condition that an object must be simple—is the best condition on offer for something to be an instance of OBJECT, given whatever aims we have for something to count as an instance of OBJECT.

To help answer that question, Korman summarizes one of the primary motivations for holding the nihilist view:

> [B]y eliminating ordinary objects, one escapes commitment to arbitrariness, coincident objects, overdetermination, overpopulation, and so on, in one fell swoop. No objects, no problems….It also draws support from the overdetermination argument: composites ought to be eliminated because, were there such things, they would be in causal competition with their atomic parts and would therefore be causally redundant. (Korman 2015:19; 20)

For example, if a bunch of *atoms arranged baseball-wise* broke a window, then the claim that *the composite baseball* broke the window would be causally redundant. For nihilism, the composite object is not needed to explain the broken window

when we already have the atoms to explain the event. Note also the normative "ought" language Korman uses that illustrates one of the conceptual ethics considerations for OBJECT: we *should* eliminate composite objects from our ontology if we seek to avoid causal redundancy.

Nihilism also avoids the charge of arbitrariness that conservatism faces. For example, it seems arbitrary to allow that a composite object like a bikini exists, with its two spatially separate parts, while not allowing that a composite object like a trog exists, with its two spatially separate parts (a given tree and a given dog). In cases like these it looks as if the conditions for OBJECT (whatever they are) fail for conservatism, because as we saw in the first chapter conservatism counts a bikini as an instance, yet for reasons that are unclear the theory denies the existence of extraordinary objects like trogs. For nihilists, denying that either exist cuts through that problem.

But the concept of OBJECT that nihilism proposes faces a host of problems, some of which were mentioned in the first chapter, and many of which seem too costly, despite some of the benefits of nihilism mentioned above. If the only objects that exist are individual atoms, then all object terms that are meant to refer to composite objects actually fail to refer. But we use terms that are meant to refer to composite objects all the time, including terms for so-called natural objects like trees and rocks, terms that aim to refer to

artifacts like tables and chairs, and terms for social objects like money and institutions. Under nihilism, none of those composite objects satisfy the conditions for OBJECT. And it doesn't help to talk about social objects as atoms arranged in a particular way. Again, as Cotnoir and Varzi observe, this move merely trades the Special Composition Question for the Special Arrangement Question: "given any composite-object sortal K, under what circumstances are there things arranged K-wise?" (Cotnoir and Varzi 2021: 183) Nihilism is as incompatible to an ontology of the (composite) social world as to an ontology of any other kind of composite objects.

Now it may somehow turn out that nihilism is *true* and gives the right conditions for OBJECT. But it isn't obvious what mind-independent facts would determine such a thing, and the concern here is to examine which object theory is most compatible with social objects, on the assumption that there are such objects. As I mentioned, many social objects like money and groups of people are highly valued, so we are motivated to find an object theory that is most compatible with affirming that there are those objects, and nihilism does not affirm such a thing. So I turn now to see how compatible conservatism is to an ontology of social objects.

III.3.2 Conditions for OBJECT Under Conservatism

Korman summarizes conservatism this way:

> Conservative views are views on which there are such ordinary objects as tables, dogs, and tree trunks but no such extraordinary objects as trogs ["an object composed of the dog and the tree trunk"], incars ["it is impossible for an incar to leave a garage"], and snowdiscalls ["an object that is made of snow, that has any shape between being round and being disc-shaped, *and* that has the following strange persistence conditions: it can survive taking on all and only shapes in that range."]. Accordingly, conservatism is a view (only) about which objects there are and aren't and is neutral on a wide variety of other questions about objects. (Korman 2015: 23)

Unlike nihilism, conservatism accepts the existence of composite objects like trees and rocks, and artifacts like tables, as long as the composite object in question is ordinary in some way. Whether some social objects count as "ordinary" can remain an open question but, contrary to nihilism, there is no principled reason under conservatism for denying the existence of all objects in the social world. But Korman himself admits that his arguments for conservatism are dialectically question-begging, meaning that he presents arguments against nihilism and permissivism while using a premise that those who affirm such theories explicitly deny, namely, that conservatism is true. We

don't get much of an argument from Korman for any *principled* way of figuring out which objects count as ordinary and which count as extraordinary (and therefore do not exist, according to conservatism). In other words, the grounding conditions are far from clear for something to count as an instance of OBJECT under conservatism; what feature or features tracks something as "ordinary" is far from clear. If under conservatism both organic objects like trees and artifacts like cars exist (among other ordinary objects), it is difficult to see what the conditions would be that allow for all those very different kinds of objects.

Though Korman fails to specify conditions for what he thinks are ordinary objects, there have been other attempts to specify grounding conditions for what counts as an ordinary or common sense object. Cotnoir and Varzi describe a few conservatist proposals that attempt to specify conditions for when such an object exists. For example, "[a]ccording to *regionalism*, it is spatial considerations that play a crucial determinative role". (Cotnoir and Varzi 2021: 188) On this view, something counts as an instance of OBJECT when it fulfills conditions that have to do with spatial features. For other conservatist views, "composition obtains instead when the objects are *functionally* united or rigidly *bonded*". Others give different conditions for OBJECT, like exhibiting "internal organizational coherence or the disposition to cause us to perceive a unity,

respectively. Still others have suggested that the common-sense conception is broadly *teleological*, so that only things that jointly serve a purpose would be said to compose."

But all of these attempts to identify some condition that is common to all and only instances of OBJECT fail in the familiar way that Wetzel observed regarding species' and words. Cotnoir and Varzi capture one of the main problems with conservatist proposals this way:

> …various candidate analyses tend to fail to problems of generality. For any purported condition ψ_φ, one can often generate counterexamples where the φs that satisfy ψ_φ do not intuitively compose anything, or where some φs compose something without satisfying ψ_φ. (Cotnoir and Varzi 2021: 188)

However we engineer the concept of OBJECT using conservative proposals, the conditions for such a concept don't seem to generate all and only instances that conservatives are after. It appears there are no clear conditions for OBJECT that conservatives agree on or that are immune to counterexamples. For these reasons it looks like the concept OBJECT under conservatism is defective in at least some ways.

Though conservatism does better than nihilism in some ways regarding its compatibility with social objects, there are two problems with the view with

respect to an ontology of the social world. First, if there are no clear, stable, predictable grounding conditions we can assess for whether some object in question would count as *ordinary* under conservatism, then there are no clear, stable, predictable grounding conditions we can assess for whether some *social* object in question would count as ordinary. Second, aside from the lack of principled conditions for what counts as ordinary, some of the examples conservatives give of extraordinary objects could turn out to be social under the right circumstances. In those cases there would be social objects that are extraordinary for conservatives, and therefore under that view would not exist. Conservatism denies that objects like trogs exist, but it is easy to imagine a scenario where something like a trog has social significance and intuitively should count as a social object, yet does not exist under a conservatist view. Suppose there is some game that develops within a community where the person who finds the most animals near a tree trunk wins: if you find a squirrel near a tree it's called a "trel" and worth 5 points, but if you find a dog next to a tree then those two things together are called a "trog" and worth 10 points. (I'm not saying it would be a *fun* game.) We could imagine other similar cases. The point is that some object that counts as "extraordinary" today under conservatism might count as a valuable *social* object tomorrow, and we will then be stuck with choosing either to affirm the value and existence of such a social

object or deny such an object exists because of some unclear, unprincipled conditions for OBJECT under conservatism.

The observation Wetzel makes that was mentioned in the first chapter, where there is no feature common to all and only tokens of a type, holds for OBJECT as well under conservatism. And if Cotnoir and Varzi are correct in the quote above, it looks as if the proposals for conditions fail in some way, because of some generally unclear restrictions for what counts as an instance of OBJECT under conservatism. As I have mentioned, if many social objects have value, then there will be normative considerations for OBJECT that motivate us to find an object theory that allows for the social objects we have today (e.g. money) and, perhaps contrary to conservatism, in principle allows other social objects that may develop later (e.g. trogs in a game). In the next section I consider whether the conditions for OBJECT under permissivism satisfy those normative goals for OBJECT, allowing for all kinds of (actual and potential) social objects. I argue that they do.

III.3.3 Conditions for OBJECT Under Permissivism

For nihilism, only atoms are instances of OBJECT. Conservatism allows for composite objects, including some social objects, so in that way is more compatible with social ontology. But it isn't clear whether the grounding conditions for something to count as an instance of OBJECT under conservatism

might rule out some valuable social objects, because it isn't clear what the grounding conditions are for something to count as an instance of OBJECT under conservatism. It would be better if there was an object theory that *in principle* allowed for *any* potential social object.

I claim that permissivism offers the benefits we want for an object theory that is most compatible with an ontology of the social world. Fairchild and Hawthorne describe permissivism this way:

> We might mention in passing two parameters along which one might be permissive. One concerns composition – the permissivist about composition, the 'universalist', will think that any plurality of material objects compose a further object. Another concerns the abundance of coincidence – here, the permissivist allows for a multitude of coincident material objects whenever there is any material object. This kind of permissivist standardly holds that the variety of coincident object have divergent, often radically divergent, essences. (Fairchild and Hawthorne 2018: 46)

The kind of permissivist described above is permissive in two ways. First, the universalist permissivist gives no restrictions for whether two or more things *compose* an object. Second, the "manyist" permissivist gives no restrictions on

how many coincident objects can be *constituted*. So for some arrangement of atoms $a_1, \ldots, a_n$ there can coincidentally be, for example, those atoms can constitute an object that is a lump and they can constitute a different object that is a statue. I will focus more on the former kind of permissivism involving composite objects, given the extensive literature on mereological composition that includes precise axioms, theorems, and definitions, compared to the lack of those sorts of elements in the literature on the constitution relation. So when I mention "permissivism" below I have in mind compositional permissivism.

Because of the lack of restrictions on the conditions for what counts as an instance of OBJECT under permissivism, permissivism offers at least two benefits relative to our current discussion on social objects. First, permissivism gets us as close as possible to describing an ontology that is *agent-independent*, i.e. an ontology that exists independently from any activity of social agents. Under permissivism (and assuming atomism), atoms are "out there" to be discovered, regardless of social activity. But unlike nihilism, which allows *only* for atoms, permissivism allows for composite objects as well, so in principle allows for any social objects that are composite objects. Under permissivism, all composite objects are also "out there" to be discovered, regardless of social

activity and, unlike conservatism, permissivism in principle allows for social objects, like trogs in a game.*

Second, the normative considerations mentioned above involving conceptual ethics for OBJECT — namely, that an object theory should allow for valuable social objects — are satisfied under permissivism. Under permissivism, any group of people composes an object (and in some cases more than one object, like in an imagined case where the Miami Dolphins and the Miami Dolphins Book Club are both composed of the same members). Monetary items like dollar bills are objects. And trogs are objects, whether some social community makes use of them now, at some later time, or never.

There may of course still be concepts of social objects that are defective in some way; permissivism does not claim that for any concept, social or otherwise, there will be instances for such a concept. For example, some social concepts involving race may be defective in all kinds of ways, so that some defective concept of RACE fails normative goals for such a concept and also fails to have instances. The concept will not be defective because of some restriction on composition, but for other reasons, and permissivism may be defective in some ways for other reasons not considered here. I am not claiming that there

---

* Although per Brouwer (2022) there may be cases of *social inconsistency*, where some proposed social object is incompatible with another proposed social object. Whether permissivism allows for social inconsistency is a separate question.

are no substantial challenges for permissivism. For reasons that nihilists give, like causal redundancy, permissivism may include defective conditions for what counts as an instance of OBJECT. But the cost of that defect may be offset by the benefits permissivism offers in allowing a plenitude of valuable social objects.

III.4. Conclusion

I have given reasons for the superiority of the 2D model over the type-token model. Through its anchoring relation, the 2D model can explain the role that social agents play in unifying many things under a kind or concept. The ongoing project of CE can further clarify how social agents can engineer the instantiation conditions for many concepts, and the many normative and evaluative considerations for concepts. The instantiation conditions then determine what objects are instantiated by the concept in question. In the context of an ontology of the social world and the objects within that world that have value, permissivism looks to be the best theory on offer if one of our normative goals is to prioritize the existence of those valuable objects. With this preferred object theory in mind, in the next chapter I take a closer look at whether there is a clean division between social objects and objects that some believe are non-social.

III.5. References

Brouwer, Thomas N.P.A. (forthcoming). Social Inconsistency. Ergo.

Balaguer, Mark. (2021). *Metaphysics, Sophistry, and Illusion: Toward a Widespread Non-Factualism*. Oxford: Oxford University Press.

Cappelen, Herman and Plunkett, David. (2020). "Introduction: A Guided Tour of Conceptual Engineering and Conceptual Ethics" in *Conceptual Engineering and Conceptual Ethics*. Alexis Burgess, Herman Cappelen, and David Plunkett, eds. Oxford University Press.

Cotnoir, A.J. and Achille C. Varzi. (2021). *Mereology*. Oxford: Oxford University Press.

Dershowitz, Naomi. (2020). "Nihilism, But Not Necessarily." *Erkenntnis*, September. https://doi.org/10.1007/s10670-020-00311-7.

Epstein, Brian. (2015). *The Ant Trap: Rebuilding the Foundations of the Social Sciences*. Oxford University Press.

Epstein, Brian. (2019). Anchoring versus Grounding: Reply to Schaffer. *Philosophy and Phenomenological Research* 99 (3): 768–81. https://doi.org/10.1111/phpr.12644.

Fairchild, Maegan, and John Hawthorne. (2018). "Against Conservatism in Metaphysics." *Royal Institute of Philosophy Supplements* 82 (July): 45–75. https://doi.org/10.1017/S1358246118000103.

Garnett, Stephen T. and Les Christidis. (2017) "Taxonomy Anarchy Hampers Conservation". *Nature* 546, 25–27. https://doi.org/10.1038/546025a

Korman, Daniel Z. (2015). *Objects: Nothing Out of the Ordinary*. Oxford: Oxford University Press.

Mason, Rebecca. (2020). "Against Social Kind Anti-Realism." *Metaphysics* 3 (1): 55–67. https://doi.org/10.5334/met.30.

Oliphint, Jared S. (2022). Using a two-dimensional model from social ontology to explain the puzzling metaphysical features of words. *Synthese* 200 (3):1-10.

Wedgwood, Ralph. (manuscript).

Wetzel, Linda. (2009). *Types and Tokens: On Abstract Objects*. MIT Press.

CHAPTER IV

SOCIAL OBJECTS AND THE METAPHYSICS OF WORDS*

Up to this point, I have been talking about *social* objects without explicitly considering whether there is a distinction between social objects and non-social objects. There seem to be intuitively clear examples of social objects. For example, a dollar bill isn't something you would find produced in a natural setting where no human society has ever existed. The object depends on social agents. Organized groups like the US Supreme Court are thought of as paradigmatic social objects for similar reasons. These objects are quite different from rocks and trees that, intuitively, are what they are independent of any social activity or influence. Some might conclude that we are motivated then to think of some unified class of objects as *social*, and to think of an entirely distinct class of objects as purely *non-social* objects.

Despite whatever initial appeal such a distinction might hold, this chapter argues against a clean division between social objects and non-social objects. If Payton (2022) is right, "the distinction between what is social and

---

* Reproduced with permission from "Springer Nature"; and "Using a Two-Dimensional Model from Social Ontology to Explain the Puzzling Metaphysical Features of Words" by Jared S. Oliphint, 2022. *Synthese*, 200: 227, Copyright 2022 by Jared S. Oliphint.

what is not rarely receives an argument". In fact, Payton goes on to site several figures working in the field of social ontology who are skeptical of whether drawing a distinction between the social and the non-social can be done:

> In particular, social ontologists have previously expressed serious doubts about the prospects of drawing the social/non-social distinction in a principled way. For example, Brian Epstein states that he is "not confident" that distinguishing clearly between social and non-social facts can be done (Epstein 2015, p. 102). Katherine Ritchie writes that "giving a noncircular definition of what it takes for something to be social is difficult, if not impossible" (Ritchie 2020, p. 404). And Sally Haslanger concurs, insofar as she thinks "it is unlikely that there is a non-circular definition [of social]" (Haslanger 2016, fn. 8). (Payton 2020: 2)

And Epstein (2021) mentions Greenwood (1997) as an example of someone who denies "that there is any criterion for distinguishing the social from the non-social".

I count myself among those who are skeptical of a disjoint distinction between the social and the non-social, and I argue that characterizing objects as social in *degrees* rather than social or non-social *simpliciter* gives us the right

characterization of objects that count as social.* Many objects we find in the natural world that we might not ordinarily characterize as social — for example, artifacts that non-human animals use — nonetheless seem to have at least some social features. If that is the case, the social theorist who seeks to propose a clean division between the social and the non-social has difficult choices to make: does she characterize an object as social if it has only *one* social feature? Or does she characterize an object as social only if it has *many* social features? Despite what might be a *prima facie* motivation to characterize social objects as a completely separate class from non-social objects, I argue that a disjoint division between social objects and non-social objects delivers several counterintuitive results. So I challenge such a characterization and then offer my own alternative view. I then turn my attention to the metaphysical features of words as objects that are social, representational devices, and argue that the role of social agents, modeled by the 2D-CE model, explains many of their puzzling features.

IV.1. Characterizing Social Objects and Non-Social Objects as Disjoint

* It will be important not to equate "social" with "socially constructed" when I mention the former. (See the aforementioned (Payton 2022) for an argument distinguishing the two.) I do not pursue the question of what makes something socially *constructed*, partially because I suspect that my main argument here—that an object can be social in *degrees*—would apply in the same way to social construction. For example, *persons* are complex beings; to say a person is exclusively either not socially constructed in any way or fully socially constructed in every way seems obviously wrong.

I take Passinsky's (2016) as a representative example of the kind of view that seeks to characterize a group of objects as social and an entirely separate group of objects as non-social. Recall from the first chapter how Passinsky (2016) characterizes social objects:

> [T]here is a unified class of "social" objects that differ in kind from "non-social" objects, and that warrant their own metaphysical investigation. Examples of social objects include borders such as the U.S.-Canada border, units of currency such as twenty dollar bills, official documents such as my U.S. passport, sovereign states such as the United States of America, international organizations such as the United Nations, and corporations such as Google. Examples of non-social objects include natural objects such as rocks, rivers, trees and persons; and ordinary material artifacts such as tables and chairs. (Passinsky 2016: 1)

According to Passinsky, then, no social object is a non-social object. The non-social objects include natural objects (though in the statement above it isn't clear whether non-social objects include *all* natural objects). All social objects are artifacts, but not all artifacts are social objects; money, for example, is social but tables and chairs are not. The condition that initially guides her classification is

whether a given object is *made to perform some social function*; if an object performs some function that "can only be performed in the context of a society", that object counts as a social object.

For Passinsky, the fact that something is a social object is grounded by the fact that it was made to perform some social function (and that fact may be grounded by other facts, as we will see). And why is *that* condition what determines what objects instantiate the concept SOCIAL OBJECT? In other words, what are the anchors that determine such a condition? That will be one of the primary questions we will be addressing in this section, doing the work alongside Passinsky of engineering the concept SOCIAL OBJECT. There are better and worse ways to engineer the concept. If social agents like Passinsky determine the conditions for something to count as an instance, we can assess the proposed conditions and look for virtues and defects. Part of what I will examine below is whether her concept SOCIAL OBJECT is defective in some way; I will argue that it is, and I will offer an alternative view.

One can, of course, simply stipulate a specific condition for an object to count as an instance of SOCIAL OBJECT. But presumably we want an account of objects that are social not by stipulation, but an account that tracks our intuitions in some way and is open at the outset to whatever kinds of objects

might end up counting as social. So that is the kind of account I will pursue and argue for.

More specifically, I argue that views like Passinsky's, where an object is either social or non-social, suffers from two problems: 1) it is vulnerable to several counterexamples, and 2) it does not give us the desired extension for what we intuitively think of as objects that are social; that is, such views do not track common usage when we speak of things as "social". Otherwise put in terms of CE, Passinsky's concept SOCIAL OBJECT is defective for those reasons. I will first cover some counterexamples to Passinsky's characterization of what counts as a social object, and then I will argue for my own view that characterizes objects as social based on whether an object exhibits social features (to be explained more fully below).

IV.2. Challenges to a Disjoint Division Between Social Objects and Non-Social Objects

IV.2.1 Artifacts

Passinsky's initial condition for an object being social is whether the object was made to perform some function within a society. So on this account, the *intent* involved in the production of an object determines whether or not the object is social. She gives a few examples:

> [T]he function that chairs are made to serve is not a social function, since something can perform the function of supporting a person in a sitting position outside the context of a society. Robinson Crusoe living alone on an island could make a chair and sit on it, and the chair would thereby perform the relevant function…Nor do artifacts such as diaries, strollers, and board games count as social objects, since the functions they are made to serve could be performed outside the context of a society. (Passinsky 2016: 7;8)

The way she applies her characterization here of what makes an object social is confusing in light of the condition she has given regarding intent. If we restate the example above in terms of the grounding relation, then the fact that chairs, board games, etc. are *non-social* objects is grounded by the fact that a chair *can* perform its function outside the context of a society. On this view, the mere capacity for performing a function outside the context of a society is something that social objects cannot do. So a further condition for something to be a social object is that it does not have the capacity for performing its function outside the context of a society. But lacking the mere capacity for performing a function outside the context of a society is a different condition for something to be a social object than the first condition she explicitly stated: the first condition

involves *intentions* when producing an object, but the condition in the quote above involves an object's *capacity* to perform its function. So on this account, if some object at t_1 was intentionally made to perform a function within a society, but there is the possibility that at t_n that object could perform its function outside the context of a society, then presumably that object would not count as a social object.

We now have two grounding conditions on Passinsky's view for something to count as a social object: 1) the fact that an object was made with the intent to perform some function within a society, and 2) the fact that an object does not have the capacity to perform that intentional function outside the context of a society. The second condition depends on the first, since it refers to the intentional function included in first condition.

Now consider a scenario involving shell money, where sea shells are currency for some community.* In a hypothetical scenario that Passinsky herself introduces, she uses the example of sea shells as currency, and indicates in the example that the sea shells would count as social objects. (Passinsky 2106: 14) Suppose at t_1 that a sea shell is a natural object and, hence, on Passinsky's account, is not a social object at that time. Suppose now that, at t_2, that sea shell

* Thanks to Nathan Howard for pointing out this example independently from it being mentioned in Passinsky (2016: 14).

is used as currency, which Passinsky has identified as a paradigmatic social kind. Hence, the sea shell is a social object at t_2. But, as a natural object, the sea shell was not made with the intent for it to perform a function involving currency within a society. So it fails to fulfill the first condition we saw above. Yet according to Passinsky, the sea shell can still count as a social object. It isn't clear why such an object can still count as social on Passinsky's account, but I assume if pressed she might say it is a social object because at some time it fulfills the second condition: the sea shell does not have the capacity to perform its currency function outside the context of a society. But we saw above that satisfying the second condition depends on satisfying the first condition, which includes the intentions involved in the object's production. Contrary to Passinsky's use of this example as a social object, if we apply the grounding conditions she herself gives for something to be a social object, the sea shells fail to satisfy her own conditions, even when the shells are used as social currency. And that seems wrong, because money is supposed to be an obvious example of a social object, as she recognizes by using such an example. So her conditions for SOCIAL OBJECT fail to be instantiated for an obvious example of a social object: sea shells as currency. Her concept looks to be defective in the way it handles paradigmatic artifacts.

IV.2.2 Persons

There are also other important objects that are not artifacts but that we think of as obviously social, namely persons, both individually in the case of one person and collectively in the case of pluralities of persons (that I will call "groups" here for convenience). Groups of persons are objects that continue to be a main subject of study within social ontology.[*] Not only do groups make up at least part of whatever it is to be a society in the first place, many groups — the US Senate, for example — fulfill one of Passinsky's conditions, viz., being designed to function in a society. Groups like the Senate also fulfill her second condition: many governmental bodies do not have the capacity to perform their intended function outside the context of a society. If the U.S. Supreme Court justices were traveling by boat and ended up stranded on a desert island, they could not perform their function as the *U.S.* Supreme Court. So at least some persons should count as social in some respect, even on Passinsky's own conditions, insofar as individual persons are members of all kinds of social groups involving family, ethnicity, and so forth. But we saw before that she classifies persons as natural, and therefore non-social. So again it looks like her concept of SOCIAL OBJECT is defective even for her own view: persons end up fulfilling

[*] See (Fine 2020), (Wilhelm 2020), and (Horden and López De Sa 2020) for just a few representative examples, as well as (Hawley 2017), cited below.

the grounding conditions for SOCIAL OBJECT, despite her claim that persons are natural, non-social objects.

IV.2.3. Non-Human Animals and Social Objects

There are also good reasons to consider some objects within the context of non-human animals as social, which could prove troublesome for a view that sharply distinguishes between social objects and natural objects. Artifacts like nests that are built by birds or other animals, for example, rocks used as "hammers" by primates, or even mounds of dirt made by termites are made from natural objects, yet many of those objects are made to perform social functions within non-human societies. An object like a termite mound does not have the capacity to perform its function—housing some termite community—outside of a termite society. So some non-human artifacts fulfill both of Passinsky's grounding conditions for being a social object, yet on her view those objects would likely be classified as natural objects on her account and therefore as non-social objects. Consider a further example: many animals are territorial, marking out boundaries for some space of land, however rough the boundaries might be. Whatever objects served as boundaries, those boundaries would be clear instances of natural objects on Passinsky's view. But on her view boundaries are also clear examples of social objects, so again we have a conflict of classification at the level of object.

Further, we could consider groups of non-human animals that are social, like a group of primates or a pride of lions. Just as in the case of human groups of persons, groups of animals also satisfy both of her grounding conditions for what counts as a social object, yet they would also count as natural, non-social objects. I mention these cases to illustrate that the scope of the social world extends broadly beyond human societies, and that objects in the natural world can exhibit social features in addition to natural features. These observations further illustrate that there is not a clean social/natural distinction.[*]

It also isn't clear on Passinsky's view whether the grounding condition requires that a particular intention must be present in the minds of only one person, two persons, or most persons involved in something like a chair's production in cases where social objects are produced by multiple persons. Suppose a fence was being constructed to act as a border between two regions. Does the intent to make a specific border need to be in the minds of all government agents who commissioned the fence, and should the same intent be in the minds of the workers who build it? Intentions in the production of objects can get quite complex and difficult to locate and track.

---

[*] Though there is some mention of the connection between social ontology and the non-human social world in (Arango 2016), for instance, the connection deserves much more treatment in the literature than it has been given here.

An implicit rejection of such a restrictive view finds support from Hawley (2017), who only briefly mentions that the way to distinguish whether an object is social is to determine whether an object has "social significance" (Ibid.: 409). Social significance as a grounding condition for an object to be social would then include at least some artifacts, but would also include any person who, as Hawley maintains, instantiates "many important properties [i.e. features] in virtue of our own attitudes and those of others". For Hawley, an object only needs social significance to count as social, so persons naturally count as social.

I believe the kind of objects Passinsky is after in her characterization is not social objects, which in the literature tends to include many different kinds of objects like groups of people and institutions. What Passinsky seems to be after is a class of what we might call *societal* objects, where societal objects fall under a much broader class of objects that are social in various ways. If one of the grounding conditions for something to count as social is that it must function in an actual, existing society, that condition would end up eliminating far too many objects that we intuitively think of as social such as groups, persons, and so forth. So there are reasons to think that the concept Passinsky is after is not SOCIAL OBJECT, because the grounding conditions she supplies for it don't seem to fit what we intuitively mean by that term. The term "SOCIETAL

OBJECT" seems like a better fit for the concept indicated by her grounding conditions, particularly because of the requirement of an actual society. However, in the next section I question whether we should classify objects as exclusively social or non-social.

IV.3. Objects As Social in Degrees

We saw above that the conditions Passinsky gives for something to be an instance of SOCIAL OBJECT makes her concept defective in two ways. First, some objects that seem intuitively social do not count as social objects on her view. Second, her conditions end up making some natural objects social, which contradicts her own classification that natural objects are non-social objects. The concept she is after seems better captured by the term "societal object". Contrary to that approach, I think there is a good case for taking objects to be social in some ways and perhaps non-social in other ways. That is, instead of characterizing some class of objects as either exclusively social or non-social, full stop, I argue that a better understanding of the social world can be achieved by taking objects to be social to varying *degrees*. Characterizing objects in this way is not vulnerable to the same counterexamples mentioned above. On this approach one, some, many, or most features of an object can be social; many features of a twenty-dollar bill, for example, are social, while perhaps fewer features of a chair, proportionately speaking, are social. Take Mount Rainier, for example. It

is intuitively a natural object and, hence, on a characterization like Passinsky's above where rocks, rivers, and trees are non-social objects, Mount Rainier would therefore also be a non-social object. Though it was not made by social agents, let alone made with any intent for it to function in a society, it clearly plays a few roles within a society: it plays an aesthetic role for those who are able to admire it from miles away, it is regularly climbed by mountain climbers for sport and leisure, and so forth. So although it is intuitively a natural object, it exhibits aesthetic social features, recreational features, and perhaps other social features that resist any characterization of it as non-social simpliciter. Similar considerations apply to any number of other natural objects, like rivers and canyons.

Instead of attempting to find necessary and sufficient conditions for something to be a social object simpliciter or a non-social object simpliciter, I believe it is far more fruitful to consider whether some object in question has any number of features that depend on social agents in some way. Roughly, what I have in mind is similar to Hawley's notion mentioned above, where the grounding condition is satisfied and an object exhibits a social feature if all or part of the object bears some significance for an individual or in the context of a society. An object can be socially significant in this way regardless of any intent in its production, or whether it has the capacity to fulfill its function in other

contexts. Persons, groups, non-human artifacts, and many human-made artifacts exhibit features that are social *in some way* or in some respects. For example, a chair used in court by a judge can be made without any intent for it to function within a society. But that chair can be used to perform an important function of signaling when the court is officially in session, for example, within a community or society in the context of a courtroom, regardless of whether such intent is present in the chair's production.

Or consider an example of a social object that exhibits both social features and non-social features, taken from Passinsky's characterization above: the sovereign state of the United States of America. There may be non-social features of the USA, like the amount of annual rainfall within its borders, or its highest point of elevation. The USA is not cleanly a social object simpliciter or a non-social object simpliciter; its many features will include a mix of social features and non-social features. This example generalizes for many objects that will exhibit a mix of social features and non-social features, motivating an account of the social world that resists an all-or-nothing, object-based account. For those who believe there can be coincident objects (in the way discussed in chapter one), those numerically distinct objects may also exhibit a mix of social features and non-social features. If there is the Paper and the Dollar, the Paper exhibits *artifactual*, social features but lacks the *currency* features if taken to a

remote island, for example. Perhaps both exhibit a non-social feature like the arrangement of atoms. Or take the Lump and the Statue; the Statue will include artifactual features just like the Paper, and the Lump may lack some of the *aesthetic* features that the Statue exhibits.* In any case, whatever numerically distinct yet coincident objects there are, for many objects they will not be exclusively social or exclusively non-social.

Persons, groups, chairs, and even many natural objects may exhibit features that are social in some way, and in those cases offer challenges to views that automatically count such things as exclusively non-social.† A vast range of objects, including many artifacts and natural objects, enjoy a mix of features that are social and some that may be non-social. As we saw in the previous chapter, under permissivism there are natural yet "extraordinary" objects like a trog that could function within society in the context of a game. For these reasons I think we should reject Passinsky's grounding conditions for what makes an object social (though perhaps accepting those grounding conditions for what makes

---

* Thanks to Robert Garcia for raising this objection.

† Though I am not here addressing whether abstract entities like social features *themselves* can be social in some respect, it is worth asking whether an account of social entities more broadly may include some abstract entities and not others. For example, if the number two is an abstract entity, we use it for all kinds of social functions like scores in sports, recipes, and so forth. If that's the case, we use social concepts, properties, and other abstract entities to describe objects, so that those entities are social in that way. I thank Nathan Howard for raising this regress problem for social features.

something societal), and further reject similar proposals for a restrictive, disjoint division between the social world and the non-social world.

Now that we have a better idea of how to go about examining the social world, in the next section I apply much of what has been said and give an account of the metaphysics of one of the most significant kinds of objects in the social world: *words*. As I mentioned in previous chapters, words make up at least some of the representational devices used in CE, and recall that Cappelen and Plunkett specifically ask what words as representational devices are. As a start, they are objects in the social world, exhibiting puzzling features that may benefit from Epstein's 2D model.

IV.4. The Metaphysical Features of Words

In March of 2017, the Lego company posted a tweet with the following statement in response to a poll about the plural of "Lego": "LEGO is always an adjective. So LEGO bricks, LEGO elements, LEGO sets, etc. Never, ever 'legos'". More recently on the show *Lego Masters*, host Will Arnett mentioned that "legos" is not a word, followed by applause from the contestants. In the LEGO community, it is a source of frustration for some when someone says "legos". If we take the claim at face value that "legos" is not a word, then what is "legos", if not a word? More generally, what are words, and what makes something a word?

Words are useful objects for the overall project here because they are *social* objects and *representational devices* used in conceptual engineering and conceptual ethics. They are ordinarily thought of as *social* objects because they are produced and used by social groups, and one of the aims of social ontology is to analyze and explain the features of such objects. As we saw in the previous chapter, social objects are not classified primarily (or perhaps at all) by their physical or intrinsic features, but primarily by the way they are used or valued by a social group. Features of a dollar bill, for example, include the way it is used or valued, in contrast to features of a tree that are more intrinsic, like its physical parts and organic functions. As we will see, words exhibit social features that are more puzzling than the natural features we find in ordinary objects like trees.

Words also play a key role in the project of CE. Consider the following example: in 2012, a handful of thinkers convened in Sydney, Australia with the purpose of creating a new social object—a word—that could represent the concept of ignoring someone in favor of one's phone (Macquarie 2014). The group had observed the pervasive (and rude) social phenomenon of ignoring others while being distracted by one's phone, and the thought was that giving this type of act a singular name would enable people to identify it more readily when it was occurring. They hoped that this would increase awareness of the

phenomenon and thereby help to decrease its occurrence. After a few candidates were considered, the group landed on the word "phubbing", combining the words "phone" and "snubbing", to represent the phenomenon in question.

Here we have an unusually clear understanding of the origins for this new social object, "phubbing": the reasons for its origin, the location of its origin, even the exact time of its origin. The way this word originated is not typically the way words develop within ordinary language, of course, but the elements of its origin are helpful for illustrative purposes, in contrast to other cases where the etymology of a word can span centuries across wide swaths of the globe.

So we could ask what kind of objects things like "phubbing" are and how they are related to concepts and other representational devices. Suppose Kylah is playing a game on her phone while Senna and Kase are attempting to talk to her. We have an instance of the concept PHUBBING. Suppose further that Clara is checking a social media feed while Seren asks her a question. We have another instance of PHUBBING. But what exactly is the *word* "phubbing", i.e. what kind of thing is it and what are its features and characteristics? (I will distinguish when I am speaking about the *concept* PHUBBING, which has instances like those I just mentioned, by continuing to represent concepts in all caps, and when I am speaking about a *word* like "phubbing" I will use quotation marks.)

There are many kinds of objects that can count as instances of a given word. Because of this, explaining the metaphysical features of words is a difficult task. The objects we call words vary far more widely in their metaphysical profile than other kinds of objects: sound waves can be words, ink markings can be words, neon lights can be words, and so forth. The degree of metaphysical variation that is allowed for words makes it difficult to give a metaphysical account for 1) what distinguishes words from non-words, and 2) what distinguishes a particular word like "color" from a different (or the same?) word like "colour". The variation between all those different kinds of objects that we can consider words is greater than the variation we allow for other kinds of objects. For example, there are many different kinds of objects that are tables, but the difference between those objects we consider tables is not as great as the difference between a sound wave we consider to be an instance of the word "tree" and an ink marking that we also consider to be an instance of the word "tree". To explain the unique, metaphysical complexity of words, I use the familiar 2D-CE model. Recall from the first chapter some of Cappelen and Plunkett's questions:

> What are the objects being assessed and improved (and do they exist)? If the aim of conceptual engineering is to assess and improve concepts…then we are ultimately on the hook for an

> account of what these objects are. (Cappelen and Plunkett 2020: 7)

And "How important is it to have a correct descriptive account of those devices in order to do the engineering project well?" (Cappelen and Plunkett 2020:14) One of the goals of this chapter is to go some way towards answering those questions.

Before applying the 2D-CE model, however, I clarify the kind of data that generate puzzles about the metaphysics of words. Any answer to questions like "What is a word?" and "What makes something a word?" must account for how the data below and the puzzles they generate are anticipated and are consistent within a working theory. I then apply the model and give an account of the metaphysical features of words as objects in the social world.

IV.4.1. The Data

Suppose I am giving directions to someone and instead of saying "Make a right at the next light" as I intend, I am suddenly overtaken by a coughing fit as I start to say the last word. The word "light" is replaced by some unrecognizable, breathy noise; is that noise a word? Recall from the first chapter Kaplan's (1991) discussion of situations where someone intends to produce some particular word, but whatever object is produced is significantly different from the

standard inscription or utterance of the intended word. The object that is produced could be another word, as in a Freudian slip, or it could be something unrecognizable like a coughing sound. This question is not merely an abstract thought experiment, but is ethically charged for those with a disability. A hearing or speech disability can cause significant deviation from the standard form of an utterance at the input (hearing) stage or output (speech) stage. Degrees of visual impairment can cause analogous issues for inscriptions. Situations like these prompt Kaplan to reject resemblance conditions for membership of word kinds. Instead, the *intent* of the person producing the word determines which word is produced.[*] For Kaplan, intent plays a greater role in explaining the metaphysical features of words than does the resemblance of objects of a certain word kind.

Recall Cappelen's (1999) response to Kaplan. He supposes that we are asked to produce a token of "l", but instead we produce an unrecognizable, squiggly shape. If there is no resemblance condition for kind membership, and

---

[*] There is some discussion in Cappelen (1999), Hawthorne and Lepore (2011), and elsewhere over exactly how intentions function within Kaplan's view. Kaplan addresses this issue in Kaplan (2011), but whatever Kaplan's actual view, let the above description of Kaplan's view represent the view where proper intentions (wholly or partially) determine kind membership for a word.

my intent is to produce "l", if intentions determine what word kind an object is an instance of, then the squiggly shape counts as "l". But that seems wrong. Cappelen believes that regardless of my intention, a squiggly shape should not count as "l", otherwise a complete lack of resemblance would allow almost any inscription to count as any other inscription. But surely intentions play *some* role in explanation, otherwise we end up with undesirable implications for those with disabilities.

Other examples remove intentions altogether from their role in word production. We are further asked by Cappelen (1999: 95) to suppose we find a piece of paper on the street with the following writing:

CAN YOU SPARE A QUARTER?

Although we don't know its source, we can use the words to ask others for a quarter. But now suppose that against all odds the writing was produced from accidental spilled ink, so that no intention was involved in its production. Can we still count the ink marks as words? Would it matter if a person was the cause of the ink spill, but that it was an accident and, hence, it was done without intent? Still other examples remove agents altogether from their role in word production. Consider an example from Epstein (2009): suppose a wave was responsible for making a mark of "Aristotle" on the beach. Is that mark a word, and what would explain an answer either way? Juvshik (2020) continues the

oceanic example by supposing that driftwood in the shape of "C A T" washes up on the shore. But he further considers whether the driftwood is a misspelling of "A C T". Can such a natural object, uncaused by human activity and absent of any intent, be misspelled?

So we have several data and several questions prompted by the examples above:

- An unrecognizable, squiggly shape is produced instead of an intended "l", perhaps because of a disability or simply by accident: do words that are unrecognizable from the intended word still count as the intended word?
- "Aristotle" or "C A T" washed up on the shore: do words that are a result of natural forces outside of any intentional agency count as words? Could such natural objects be an instance of misspelling?
- The source that produced CAN YOU SPARE A QUARTER? is either unknown or accidental: does that affect whether the parts of the inscription count as words?

The data get more complex when we consider a word's relation to its meaning. Gasparri (2019) gives a few examples that illustrate further puzzles:

- *Change*. Words change in meaning. In one type of shift, called "narrowing" (Bloomfield 1933), the meaning of the

> word undergoes a change from a superordinate conceptual level to a subordinate conceptual level. For example, SKYLINE originally designated the horizon as such, but it is now used to refer to the horizon qua populated by a particular kind of urban landscape.

- *Homonymy and polysemy*. Consider "fluke." It can designate certain types of flatfish, the end parts of an anchor, either half of the triangular tail of a whale, and a stroke of luck. The fourth meaning is distant from the previous three, so it plausibly deserves a dedicated count, say, FLUKE_4. But what about the other three? Can they be associated to a single polysemous word, or should they be understood as senses of numerically distinct words?

In some cases, as illustrated with "skyline", it appears that one and the same word can change meaning, and in other cases, as illustrated with 'fluke', one and the same string might in fact be associated with distinct words. If a theory of the metaphysical features of words seeks to be maximally thorough, it will consider such examples that illustrate some of the most common features that words exhibit.

IV.4.2. The 2D-CE Model Applied to Words

Recall from previous chapters that the one-dimensional type-token model concludes from the lack of resemblance between some objects of the same type that kind membership must be brute or fundamental. The 2D model allows for anchors that explain the lack of resemblance conditions for objects to fall under the same kind, due to the way social agents have set up kind membership for

objects. For words, within particular media-kinds like *inscriptions* or *utterances*, resemblance conditions for kind membership may be put in place by social agents. An inscription like "tree", for example, allows for a good deal of variation between objects of that same inscription kind, but there needs to be *some* resemblance between those objects, which explains why an inscription like "free" does not fall under the same inscription kind as "tree".

But for words that allow inscriptions *and* utterances, and perhaps other sorts of objects as well, to count as instances of the same word, we will not find resemblance conditions across all those various objects. For instance, a word like "tree" can have instances from such diverse kinds as sign language, Braille, Morse code, and so forth.[*] A *sign gesture* that is an instance of "tree" bears no relevant resemblance to *Braille bumps* that are also instances of "tree". In place of resemblance conditions for words like "tree", there will be some complex,

---

[*] Cappellen (1999), 99; Wetzel (2009), 61; Epstein (2009), 57, for example. As one anonymous reviewer pointed out, there may be different *senses* of "word" that each correspond to a particular object type. One sense of "word", for example, could correspond to inscriptions, another to utterances, and so forth. If that's the case, there are anchoring conditions for "word" that we could explore and are worth developing, but would sidetrack us too much from the more general metaphysical picture here.

linguistic mapping system that social agents set up so that an inscription of "tree" will be an instance of the word "tree" no less than an utterance of it is.

The 2D model shows us where to look for explanations of the data we considered: we look both to the objects and to the anchors established by social agents, relative to some social community. Is the driftwood that washed up on the shore in the shape of "Aristotle" a word? In a sense it depends on *which shore* we are talking about. We are asking whether something counts as a social object, so the answer will be relative to various historically contingent, sometimes unpredictable decisions made by a social community. To take a hypothetical example, the driftwood would not be a word for a social community that is only familiar with hieroglyphic inscriptions, and thus has given no membership conditions for an object like "Aristotle". But the driftwood will be a word for a social community that has given membership conditions for objects shaped like "Aristotle". Asking whether some natural object is either this or that word simpliciter, without respect to some social community, fails to recognize the social features of words. Those social features involving social agents' anchoring are explained in the 2D model.

Likewise for the cases of homophony and polysemy such as the different instances of "fluke" in section one. These cases illustrate what initially appears to be a strange metaphysical feature of words, where qualitatively identical

objects are instances of different words. If we want to explain such a phenomenon, we should be looking not to the qualitatively identical objects, but to social agents and the historically contingent conditions they give for membership of whatever word is in question. For many objects, we typically think of qualitatively identical objects as falling under the same kind or concept, like two qualitatively identical tables. But the 2D model can explain how different syntactical and semantic conditions are given by social agents for an object to fall under $FLUKE_1$ but not $FLUKE_4$, for example. There is no unique feature of an instance of "fluke" that can *ground* the fact that it falls under $FLUKE_1$ but not $FLUKE_4$. But there are unique anchoring conditions involving social agents and the syntactical and semantic conditions they give to distinguish when "fluke" falls under $FLUKE_1$ but not $FLUKE_4$. The model also explains how some word like "skyline" can satisfy different agent-dependent conditions over time. As the appearance of the earth's horizon changes for social agents because of the increasing number of skyscrapers, over time agents slowly changed the anchoring conditions (the syntactical and semantic conditions) for the word "skyline".

Words can be as unruly as the agents who produce them, and as subjects of metaphysical inquiry they can generate data and questions that admit a high degree of metaphysical complexity. I have applied work from social ontology

and CE to the ongoing discussions surrounding the metaphysics of words as representational devices, and in that way charted a new path within the literature on the metaphysics of words.

Of course more work remains, including conceptual ethics questions, like what kinds of features some particular word should admit. The question to ask then is, "*Should* we include such and such conditions for an object to be an instance of *x*?", such as conditions that would allow or exclude emojis from being words or, as we saw above, conditions that would allow for quite a bit of latitude in the appearance or sound of some object for the sake of those with a disability. Or we could ask, "Should we include such and such conditions for a device to represent an instance of PERSON?" The questions acknowledge the choices that social agents can make when giving reasons for the agent-dependent, normative conditions of words.

IV.5. References

Arango, Alejandro. (2016). "Animal Groups and Social Ontology: An Argument from the Phenomenology of Behavior." *Phenomenology and the Cognitive Sciences* 15 (3): 403–22. https://doi.org/10.1007/s11097-015-9430-2.

Brouwer, Thomas N.P.A. (forthcoming). Social Inconsistency. *Ergo.*

Cappelen, Herman. (1999). Intentions in Words. *Noûs* 33 (1): 92–102.

https://doi.org/10.1111/0029-4624.00143.

Epstein, Brian. (2009). Grounds, Convention, and the Metaphysics of Linguistic Tokens. *Croatian Journal of Philosophy* 9 (1): 45–67.

Epstein, Brian. (2015). *The Ant Trap: Rebuilding the Foundations of the Social Sciences*. Oxford University Press.

Epstein, Brian. "Social Ontology". *The Stanford Encyclopedia of Philosophy* (Winter 2021 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/win2021/entries/social-ontology/>.

Fine, Kit. (2020). "The Identity of Social Groups." *Metaphysics* 3 (1): 81–91. https://doi.org/10.5334/met.45.

Gasparri, Luca. (2020). A Pluralistic Theory of Wordhood. *Mind & Language* 2020: 1-18. https://doi.org/10.1111/mila.12297.

Greenwood, John D. (ed.). (1997). *The Mark of the Social: Discovery or Invention?*, Lanham, MD: Rowman and Littlefield.

Hawley, Katherine. (2017). "Social Mereology." *Journal of the American Philosophical Association* 3 (4): 395–411. https://doi.org/10.1017/apa.2017.33.

Horden, John, and Dan López de Sa. (2020). "Groups as Pluralities." *Synthese*, August. https://doi.org/10.1007/s11229-020-02715-y.

Juvshik, Tim. (2020). Good "Cat", Bad "Act." *Philosophia*, November. https://doi.org/10.1007/s11406-020-00287-2.

Kaplan, David. (1990). Words. *Aristotelian Society Supplementary, 64*, 93–119. https://doi.org/10.1093/aristoteliansupp/64.1.93.

Kaplan, David. (2011). Words on Words. *Journal of Philosophy*, 108 (9), 504-529.

Korman, D. Z. (2015). *Objects: Nothing Out of the Ordinary*. Oxford University Press UK.

Miller, J. T. M. (2019). A Bundle Theory of Words. *Synthese*, October. https://doi.org/10.1007/s11229-019-02430-3.

Passinsky, Asya. (2016). *The Metaphysics of Social Objects*. NYU dissertation.

Payton, Dee. (2022). "Searching for Social Properties." *Philosophy and Phenomenological Research.* https://doi.org/10.1111/phpr.12901.

Wetzel, Linda. (2009). *Types and Tokens: On Abstract Objects*. MIT Press.

Wilhelm, Isaac. (2020). "The Stage Theory of Groups." *Australasian Journal of Philosophy* 98 (4): 661–74. https://doi.org/10.1080/00048402.2019.1699587.

CHAPTER V

CONCLUSION

I began the project by asking the following questions:

1) Can we better understand some features of objects that have typically been thought of as mind-independent (composition, kind membership) if we think of those features as dependent on social agents in some way?
2) What distinguishes social objects from other kinds of objects?
3) If objects used in language (like words) are social objects, in what ways do the features of objects depend on our use of language, if at all?
4) What normative (especially moral) considerations might there be if there are different ways for social agents to characterize a given object?

I addressed those questions by introducing the 2D-CE model and applying it to object theories, then particularly *social* objects, and then particularly *words* as social objects. Throughout this project I have sought to show the various ways in which agents play a role in determining at least some of the features of all kinds of objects. But there are more questions to address, such as whether we can speak of objects that have no social features whatsoever, and how we can explain the differences between agent-dependent facts (if there are any) and agent-independent facts.

Future work will expand the scope of how agents determine features of other kinds of entities. In this dissertation, I have focused primarily on material objects. But I am interested in the role that social agents play in determining the features of abstract entities as well. This project stands in a tradition within metaphysics that understands entities like kinds, concepts, and types as universal-like entities that have instances. But a closer examination that asks what the differences are, if any, between these entities seems like a worthwhile expansion of the project.

I also have in mind work that takes the 2D-CE model and applies it to *technological* or *digital* objects like NFT's, cryptocurrencies, and other digital objects of value. For example, Wildman and McDonnell (2020) are committed to the view that digital objects do not exist, yet they recognize that these allegedly non-existent objects can be stolen, and hence are forced to resolve that tension. There may be a better way to think about the character of digital objects as artifacts whose existence conditions depend on agents in some way. Apart from these applications, the 2D-CE model is intended to be a metaphilosophical tool that can be applied to many different topics, so the observations here will likely show up implicitly throughout future work.

V.1. References

Wildman, Nathan, and Neil McDonnell. (2020). “The Puzzle of Virtual Theft.” *Analysis* 80 (3): 493–99. https://doi.org/10.1093/analys/anaa005.