Assert This: "Philosophers Are Engineers"

A Study of Philosophical Engineering and Generic Judgments

Mirela Fuš

A Ph.D. dissertation to be submitted to
the University of Oslo and the University of St Andrews
for the double badge degree of Doctor of Philosophy

University of Oslo
Faculty of Humanities
Department of Philosophy,
Classics, History of Art and Ideas

Primary supervisor: Prof. Herman Cappelen
Co-supervisor: Dr. Joanna Pollock

University of St Andrews
School of Philosophical,
Anthropological and Film Studies
Department of Philosophy

Primary supervisor: Dr. Patrick Greenough
Secondary supervisor: Dr. Kevin Scharp

January, 2021

Mirela Fuš

Assert This:

"Philosophers Are Engineers"

A Study of Philosophical Engineering and Generic Judgments

Assert This: "Philosophers Are Engineers"

A Study of Philosophical Engineering and Generic Judgments

To my parents

Jozefina and Željko Fuš

&

In memory of

Niko Maksić

ACKNOWLEDGMENTS

Through the course of writing this dissertation, many more places and people than I can recall have formed me intellectually and as a person. I am deeply indebted to all of you for making this journey less lonesome and more connected. My apologies in advance for any omissions.

This dissertation was written in two places primarily: Oslo and St Andrews. Apart form my offices at Blindern and Arché many of the initial ideas were brewed at two other places: Litteraturhuset and Taste. Portions of thinking and writing were also conducted at several other places including Akureyri, Amsterdam, Bergen, Bochum, Budapest, Dubrovnik, Edinburgh, Hamburg, Kotoriba, Lørenskog, Maribor, Montreal, Munich, Paris, Reykjavík, Rijeka, Sherbrooke, Stockholm, Trondheim, Uppsala, Vienna, and Zagreb.

I had the privilege to have four supervisors. Thank you to Herman Cappelen for challenging conversations and for overseeing my dissertation project from start to finish. Thank you to Joanna Pollock for invaluable critical feedback and guidance, especially during the final stages of writing this dissertation. Thank you to Patrick Greenough for generous encouragement and keen insight. Thank you to Kevin Scharp for stimulative discussions and valuable advice.

I am deeply grateful to several colleagues who provided me with verbal and/or written critical remarks on earlier drafts of the dissertation. Thank you to James Andow, Derek Ball, Delia Belleri, Paul-Mikhail Catapang Podosky, David Chalmers, Anna Daria Drozdzowicz, Yvonne Hütter-Almerigi, Manuel Gustavo Isaac, Sigurd Jorem, Steffen Koch, David Liebesman, Jennifer Nado, Bernhard Nickel, Mark Pinder, David Plunkett, Marjorie Rhodes, Sam Roberts, Alessandro Rossi, Sarah Sawyer, Giorgio Sbardolini, Mona Simion, Rachel Sterken, Joost Vecht, and Keith Wilson. Many thanks to Sigurd Jorem for translating an English summary into Norwegian and to Henning Holmedal for copyediting the Norwegian version of the dissertation summary.

I am honored for the opportunity to have served as an active member at two research centers: *ConceptLab* (Oslo) and *Arché* (St Andrews). I was also fortunate enough to be a member of: *Conceptual Engineering Virtual Network*, *Oslo Mind Group*, *Centre for the Study of Mind in Nature (Oslo)*, *MAP Chapter of St Andrews and Stirling*, *European Early Career Philosophers Network*, *Center for Language Research (Rijeka)*, *Culture of Critical Thinking Initiative (Rijeka)*, *Society for Women and Minorities in Philosophy*. A big thank you to all the members for sharing these creative and stimulative social group environments.

I am immensely grateful to have received tremendous amounts of knowledge, inspiration, and support from so many excellent professionals, colleagues and friends from all around the world. Thank you to Solveig Aasen, Farbod Akhlaghi-Ghaffarokh, Elvio Baccarini, Marvin Backes, Suchitra Badhwar,

Conrad Bakka, Lisa Bastian, Boran Berčić, Katharina Bernhard, Edith Berntzen Hesselberg, Malte Bischof, Martina Blečić, Mark Bowker, Dragana Bozin, Aksel Braanen Sterri, Zdenka Brzović, Elisabeth Cantalopemass, Bianca Cepollaro, Ivan Cerovac, François Claveau, Maria Jimena Clavel Vázquez, Tomislav Čop, Josh Dever, Michael Devitt, Hege Dypedokk Johnsen, Anne Eaton, Miguel Egler, Matti Eklund, Jola Feix, Claire Field, Jade Fletcher, Vera Flocke, Yael Friedman, Peter Fritz, Ludvig Fæhn Fuglestvedt, Joaquim Giannotti, Ephraim Glick, Barry Guihen, Stian Gullner Klasbu, Josh Habgood-Coote, Dan Healey, Katharina Heinz, Samia Hesni, Jakob Hinze, Joachim Horvath, Savvas Ioannou, Viktor Ivanković, Bruno Jacinto, Jonathan Jenkins Ichikawa, Markus Jerkø, Robin Jeshion, Marko Jurjako, Dunja Jutronić, Chris Kelp, Alison Kerr, Kris Kersa, Helene Killmer, Dirk Kindermann, Max Johannes Kippersund, Friderik Klampfer, Nikolai Klix, Kirun Kumar Sankaran, Ethan Landes, Lorenzo Lazzarini, Øystein Linnebo, Annina Loets, Guido Löhr, Luca Malatesti, Poppy Mankowitz, Deborah Marber, Christopher Masterman, Matthew McKeever, Matthew McGrath, Nenad Miščević, Ole Martin Moen, Annikken Moer, Hannah Monsrud Sandvik, Hanne Sofie Nillson, Anh-Quan Nguyen, Andrew Peet, Walter Pedriali, Quentin Pharr, Kim Phillips Pedersen, Snježana Prijić-Samaržija, Simon Prosser, Martina Rajnović, Oda Victoria Reitan, Jeroen Rijnders, Monica Roland, Martina Rosola, Lewis Ross, Nawar Saddik, Boris Savić, Henry Schiller, Maria Seim, Camilla Serck-Hanssen, Feroz Mehmed Shah, Wes Skolits, Joe Slater, Anna Smajdor, Nenad Smokrović, Hans Robin Solberg, Sergiu Spătan, Saranga Sudarshan, Kristoffer Sundberg, Kristina Škaro, Predrag Šustar, Fenner Tanswell, Ravi Thakral, Clotilde Torregrossa, Ingvild Torsen, Caroline Torpe Touborg, Leo Townsend, Tolgahan Toy, Majda Trobok, Sascha Troscheit, Iris Vidmar Jovanović, Sara Vikesdal, Sebastian Watzl, Natalia Weights-Hickman, Robbie Williams, Timothy Williamson, Alex Yates, Alper Yavuz, Yelena Yermakova, Dan Zeman, and Berislav Žarnić.

I wish to thank to the audiences at the following events for their critical feedback: *Conceptual Engineering Online Workshop* (online, 2020), *Conceptual Engineering Virtual Network* (online, 2020), *Oslo Mind Group* (online, 2020), *ConceptLab Taking Stock* (online, 2020), *Arché 20th Anniversary Conference* (St Andrews, 2019), *2nd EXTRA workshop on Inconsistent Concepts and Conceptual Engineering* (Bochum, 2019), *Institutskolloquium* (Hamburg, 2019), *Fourth PLM Workshop: Meaning and Intention* (Vienna, 2018), *Mind, World and Action* (Dubrovnik, 2018), *Arché Work in Progress Seminar* (St Andrews, 2018), *Workshop with David Chalmers* (Vienna, 2018), *Arché Conceptual Engineering Seminar* (St Andrews, 2018), *Arché Language and Mind Seminar* (St Andrews, 2018), *St Andrews Friday Seminar* (St Andrews, 2018), *Generic Generalizations: Meaning and Social Practices* (Sherbrooke, 2018), *10th Arché Graduate Conference* (St Andrews, 2017), *Arché Conceptual Engineering Seminar* (St Andrews, 2017), *ECAP 9* (Munich, 2017), *Filosofisk fredagsseminar* (Oslo, 2017), *ConceptLab Flash Workshop* (Oslo, 2017), *Centre for Language Research*

Seminar (Rijeka, 2017), *Early Career Workshop on Relativism, Pluralism and Contextualism* (Vienna, 2017), *CSMN Work in Progress Seminar* (Oslo, 2017).

I greatly appreciate professional support from the administrative staff at the University of Oslo and the University of St Andrews. Thank you to Christine Amadou, Anne Lise Bækholt Pound, Julie Bårdsen Tøllefsen, Andrea Dale Wefring, Beate Elvebakk, Karen Haugland, Lynn Hynd, Solveig Johansen, Grethe Netland, Aina Sandvik, Caroline Strutz Skei, Pia Søndergaard, and Lina Tosterud.

I kindly acknowledge the University of Oslo and the University of St Andrews for their financial support.

I am extremely lucky to have received an extraordinary amount of support and care from several friends over the past few years. Special thanks to Joey Pollock, Sam Roberts, Lucija Ružman, Iris Segers, Antonija Tisaj Savić, Joost Vecht, and Hilde Vinje.

A thousand thanks to Henning Holmedal, for his love and devotion, and for being awesome in more than eighty-six ways, especially throughout the final stages of writing this dissertation.

Last but foremost, for their unceasing love and trust in me I am forever indebted to my family. I could have not done this without you.

This dissertation is dedicated with love to the memory of Niko Maksić.

Mirela Fuš

Oslo, 2021

CONTENTS

GENERAL INTRODUCTION

Overarching Research Question

This dissertation brings together and contributes to recent debates on *conceptual engineering* and *social group generics*.[1] It lies primarily at the intersection of philosophical methodology and philosophy of language. Moreover, it includes related areas in philosophy of mind, social ontology, evolutionary psychology, and feminist theory.

A fast-moving methodological field that most prominently goes by the name of 'conceptual engineering' (see Cappelen 2018; Burgess, Cappelen and Plunkett 2020) has recently been establishing itself as an alternative philosophical method to conceptual analysis and experimental philosophy. Under the assumption that our *concepts* or, more generally, *representational devices* can be deficient (see Scharp 2013; Burgess and Plunkett 2013a,b; Cappelen 2018), conceptual engineering projects rest on the idea that a lot of philosophical endeavor with representational devices *does not* (and, perhaps, *should not*) happen merely at a level of *description*. Instead, the advocates of engineering projects in philosophy believe that our involvement *does* (and, perhaps, *should*) occur at a level of *evaluation* and *revision*. I will refer to an engagement in these activities as a 'first order project' ('FOP'). Paradigmatic cases of conceptual engineering include FOPs about *philosophical concepts* such as BELIEF, TRUTH, and CONCEPTUAL ENGINEERING; *scientific concepts* such as WEIGHT, and SPACETIME; *social concepts* such as RACE, and GENDER; *publicly debated concepts* such as MARRIAGE, RAPE, REFUGEE, TERRORIST, and SANDWICH, etc.

Generics or *generic statements* express generalizations but do not specify how many members of the kind have the property being expressed. In many cases, generics such as "Tigers have stripes" or "Mosquitos carry the West Nile virus" are considered to be *useful* generalizations. However, a particular subclass of generics, the so-called 'social group generics' such as "Blacks are violent," "Muslims are terrorists," "Women are submissive," have recently become prominent for their morally, socially and politically pernicious effects (see Haslanger 2011; Leslie 2017; Saul 2017; Ritchie 2019). Their perniciousness has become a pressing issue since generics are ubiquitous in natural languages and most speakers use generics on an everyday basis.

The overarching research question of this dissertation is the following: *Can the pernicious effects of generics be ameliorated by applying the method of engineering in philosophy?*

---

[1] In this dissertation, I use italics to emphasize a single word or a phrase (e.g. *philosophical engineering)*. I use single quotation marks (e.g. 'philosophical engineering') to mention or introduce linguistic items. I use small caps (e.g. PHILOSOPHICAL ENGINEERING) to pick out concepts. I use double quotation marks (e.g. "philosophical engineering") for a variety of tasks, including quoting other authors, scare quotes, and mixes of use and mention.

In response, this dissertation introduces and further develops two topics: *philosophical engineering* and *generic judgments*. It addresses the above research question in two parts. Part I of this dissertation does some groundwork for the philosophical method of engineering in philosophy, its *nature* and *terminology*. Part II of this dissertation brings to the foreground *generic judgments*, their relation to *generic truths* and their role in *ameliorating pernicious effects* of generic propositions. It is, however, important to stress that, in addition to aiming at being an independent contribution to the methodological field of engineering in philosophy, Part I of this dissertation is largely in service of Part II, where generics are treated as a *case study* of the philosophical method developed in Part I.

Outline of the Dissertation

This dissertation is best understood as a monograph. It comprises a *general introduction* and *two parts*. Below is a brief overview of each chapter of the two parts of this dissertation, respectively.

Part I, "On the Nature of Philosophical Engineering: Towards Substantive and Terminological Dispute Resolution," comprises four chapters.

In Chapter 1, "Three Kinds of Disputes Behind the Unity Assumption," I argue that those who theorize about the philosophical method of FOPs submit to a common assumption about the methodology of FOPs. I take on board the following three assumptions that Cappelen and Plunkett (Cappelen and Plunkett 2020) claim are underlying FOPs (either disjunctively or in conjunction): (i) *Unifying Activity Assumption:* FOPs are instances of the same kind of activity; (ii) *Unifying Subject Matter Assumption:* FOPs share the same kind of subject matter; (iii) *Unifying Methodology Assumption:* FOPs share the same kind of philosophical methodology. I adopt a version of a *conjunction* of these assumptions which says that the *Unifying Methodology Assumption* consists in *Unifying Subject Matter Assumption* and *Unifying Activity Assumption*. I call this the 'Unity Assumption' (UA). Furthermore, I argue that despite a general background agreement about UA, there have been three kinds of disputes about UA: *merely verbal* (a disagreement concerning the *meaning* of the term(s) for the methodology behind UA), *substantive* (a disagreement concerning what the methodology behind UA *is*) and *terminological* (a disagreement concerning which *term* for the methodology behind UA *to use*).

In Chapter 2, "Substantive Dispute Resolution of the Unity Assumption," I offer a resolution of the *substantive* dispute behind UA. I argue that none of the current proposals can settle the substantive dispute about UA. The reason for that is, I argue, that the so-far proposed theories for the philosophical method behind UA cover *the unifying subject matter* and *the unifying activity scope* only partially. This further suggests that a substantive disagreement could not be resolved by settling on one of the current proposals. Instead, by arguing for a more plausible scope of the unifying subject matter and the unifying activity behind UA, I offer another way to resolve the substantive disputes over the unifying

subject matter and unifying activity that comprise the unifying methodology of FOPs. I engage in a *substantive dispute resolution* of UA as follows. First, I argue that we should consider *being philosophical* as the unifying subject matter behind UA. More specifically, I suggest that what makes the method behind UA *philosophical* is that philosophy as its *subject matter* is being engineered. This suggestion relies on the assumption that philosophy as a FOPs' subject matter can be engineered via engineering philosophical objects. Second, I argue that we should consider *engineering* as the unifying activity behind UA. More specifically, I suggest that what makes the method behind UA *engineering* is that it comprises a *five-stage* (*identification, evaluation, strategic planning, implementation, re-evaluation*) *recursive activity*. This suggestion takes engineering activity in a non-metaphorical sense and endorses engineering as a philosophical method that bears a *family resemblance* and is in *continuity with other engineering fields*. In a nutshell, *philosophical engineering* is engaging in a five-stage recursive activity that can operate on different philosophical objects.

In Chapter 3, "Terminological Dispute Resolution of the Unity Assumption," I offer a novel resolution of the *terminological* dispute about the philosophical method behind the Unity Assumption (UA). First, I introduce a general definition of a *terminological dispute* and, as a case study, I suggest a terminological dispute about the philosophical method behind UA. Second, I evaluate two prominent terminological choices, i.e. 'conceptual engineering' (see Scharp 2013, 2020; Eklund 2015; Cappelen 2018; Burgess, Cappelen and Plunkett 2020; Chalmers 2020) and 'conceptual ethics' (see Burgess and Plunkett 2013a,b; Plunkett and Sundell 2013; Burgess 2014; Burgess and Plunkett 2020), against the two main conditions that I suggest one should take on board when making terminological choices for the philosophical method behind UA: (i) *the semantic correctness condition* drawing on Belleri (2018), and (ii) *the beneficial lexical effects condition* drawing on Cappelen (2018) and Chalmers (2011, 2020). Finally, I engage in a *terminological dispute resolution* by arguing that we should introduce a new term, 'philosophical engineering.' For the philosophical method in question this terminological choice, I argue, fares better against its competition with respect to both of the above mentioned conditions.

In Chapter 4, "Objections and Replies," I address some objections concerning the substantive and the terminological dispute resolutions of the Unity Assumption (UA) that I offered in Chapter 2 and Chapter 3, respectively. In particular, with regard to the substantive dispute resolution, I discuss: Explanatory Power of Philosophical Engineering; Philosophical Engineering Outside of Philosophy; and Deutsch's Pessimistic Dilemma. With regard to the terminological dispute resolution, I discuss: Semantic Correctness or Descriptive Adequacy; Aptness to Cause Misunderstanding, Miscommunication, or Harm; Too Broad and Too Narrow Terminological Choice; What 'Philosophical' Means is Just as Unclear as What 'Conceptual' Means; and No Decisive Reasons for Introducing 'Philosophical Engineering.'

Part II, "Pernicious, Beneficial and Neutralizing Generic Judgments: A Case Study of Philosophical Engineering," comprises two chapters.

In Chapter 5, "A Darwinian Dilemma for Realist Theories of Genericity," I bring to the foreground *judgments about generic propositions*. In her paper called "A Darwinian Dilemma for Realist Theories of Value" Sharon Street (2006) argues against *realist theories of value*. In this chapter, I consider the prospects of applying an analogous Darwinian dilemma for *realist theories of genericity* by offering a debunking argument against genericity drawing on a direct analogy with Street's (2006) paper. In particular, I focus on generic statements such as "A tiger is striped" or "Muslims are terrorists" and look at the relation between *generic judgments* and *independent generic truths* expressed by generic statements posited by the realist about genericity. This chapter makes a broader methodological point about the connection between realist theories about genericity and speakers' judgments about generic propositions, based on the assumption that generic judgments have been indirectly yet significantly influenced by *evolutionary forces*.

In Chapter 6, "The Speech Act Approach to Engineering Generic Judgments," I address the following question: Can the pernicious effects of generic statements such as "Blacks are violent," "Women are submissive," or "Muslims are terrorists" be ameliorated by applying the method of engineering in philosophy? My contribution to the debate about the so-called pernicious generics is twofold: (i) I identify and critically examine current approaches to pernicious generics (Haslanger 2011; Leslie 2017; Saul 2017; Ritchie 2019) as domain-specific instances of philosophical engineering of generics; (ii) I develop an alternative domain-specific account of philosophical engineering of pernicious generics. In particular, I argue that the pernicious effects of generics stem most fundamentally from our judgments about generic propositions. As one of the plausible strategies to achieve engineering of generic judgments, I suggest the Speech Act Approach, an approach to changing doxastic propositional attitudes towards generics by utilizing speech acts. I show that the Speech Act Approach has three key benefits. Firstly, it does not depend on any particular view about the semantics, epistemology or metaphysics of generics. Secondly, it is more feasible to implement than the proposed alternatives. Thirdly, it provides a better framework for utilizing generics in order to achieve beneficial or neutralizing effects. Finally, I sidestep an ethical/epistemic dilemma for the Speech Act Approach to engineering generic judgments.

Brief Methodological Remark

In this dissertation, I mainly utilize the philosophical methods of *philosophical analysis* and *philosophical engineering*. I also endorse the existing results of *empirical research (especially from cognitive and evolutionary psychology)* where appropriate.

In Part I of this dissertation, I develop methodological tools which I utilize in Part II. In Chapter 1, I engage in *philosophical analysis* of the current state of the art of theorizing about the methodological field of engineering in philosophy. In Chapter 2, I engage in *engineering* of the philosophical *method* of philosophical engineering. In Chapter 3, I engage in *engineering* of the *term* 'philosophical engineering' for the method of philosophical engineering.

In Part II of this dissertation, in addition to *philosophical analysis* and relying on the existing results of *empirical research*, I utilize the methodological tools I develop in Part I. In Chapter 5, I apply philosophical analysis and offer a *debunking argument* against the realist theories of genericity based on the results from the *evolutionary and psychological research*. In Chapter 6, in order to develop a strategy for ameliorating the pernicious effects of generics, I develop and suggest a subkind of philosophical engineering, i.e. *pragmatic engineering* that by utilizing *speech acts* aims at changing *propositional attitudes* towards generic propositions and, consequently, generic judgments. In addition to that, I endorse utilization of *empirical research*, especially during the strategic planning and the implementation stage.

Brief Discussion

Part I of this dissertation contributes to *developing of a general framework* for the philosophical method of FOPs.

In Chapter 1, I prepare terrain for establishing the method. I elucidate the main underlying assumption about the philosophical method common to those who theorize about engineering projects in philosophy, and I identify and discuss three kinds of disputes about this assumption: *merely verbal*, *substantive* and *terminological*. In Chapter 2, I resolve the *substantive* dispute by endorsing: (i) *pluralism about philosophical objects of engineering*; (ii) *the five-stage recursive model* for this philosophical method. In Chapter 3, I resolve the *terminological* dispute by introducing a new term 'philosophical engineering' for this philosophical method.

In result, Part I contributes to the current *debate about conceptual engineering* in two main ways: (i) my proposed *substantive dispute resolution* leads to broadening of the scope and building a general model for this philosophical method; (ii) my proposed *terminological dispute resolution* leads to a better term for this philosophical method.

Part II of this dissertation contributes to *establishing pernicious generic judgments as relevant to engineering projects* in philosophy and *offers a domain-specific account for their amelioration*.

In Chapter 5, I contribute to a deeper understanding of the metaphysical nature of generics. I offer a skeptical line that suggests that our *generics judgments* may not track *generic truths* posited by the *realists about generic truths*. I provide an *evolutionary debunking argument* which offers an alternative, scientific explanation about the relationship between our generic judgments and the independent

generic truths. In Chapter 6, after elucidating a strand of literature on *pernicious generics*, I contribute to better understanding of the *source of perniciousness of generics*, and I suggest how to *change* some of our *generic judgments* by utilizing the method of philosophical engineering. In particular, in order to achieve engineering of pernicious generic judgments, I put forward The Speech Act Approach, a subkind of *pragmatic engineering of generics* as a novel, alternative domain-specific account of philosophical engineering that aims at amelioration of *pernicious* but also at utilization of *beneficial* generics. The main title of this dissertation embodies such *beneficial* utilization of generics by promoting an *assertion* of a generic proposition: "Philosophers are engineers."

In result, Part II contributes to the current *debate about pernicious generics* and supports the theoretical assumptions of Part I. Moreover, given the *engineering holism* and *division of engineering labour*[2] that I advocate, the results of Part II bear importance and potential not only for *philosophical methodology* and the *metaphysics of generics*, but also for *social reality*, by suggesting a way to contribute to amelioration of *social*, *moral*, and *political injustice*.

To sum up, this dissertation makes several important contributions, some of which may span far beyond philosophy. In particular, it establishes philosophical engineering as a *general philosophical method*. It puts forward a *domain-specific engineering project in philosophy* that takes generic propositions as FOPs. It makes a point about the relation between generic judgments and generic truths that has consequences for further understanding of the *metaphysics* and *semantics of generics*. Finally, it shows how philosophical engineering could contribute to other kinds of engineerings such as *cognitive* and *social engineering*.

Further Work

It is hard to stress enough how fast-moving the field of engineering in philosophy is. At the same time, however, we are still in the early stages of the development of this methodological field, both with respect to *general theorizing* about its foundations as well as with respect to putting forward and implementing its *domain-specific* and *particular* projects.

This dissertation theorizes about the foundations of this methodological field and, while doing so, it also proposes and engages in several engineering projects. Let me highlight three of the most prominent ones: (i) a *general project* of engineering of the *concept* (and, hopefully, the *method*) of PHILOSOPHICAL ENGINEERING; (ii) *a particular project* of engineering of the *term* 'philosophical engineering'; (iii) a *domain-specific project* of engineering of generic judgments. Nevertheless, one may argue that, in a

[2] Both between different sub-engineerings within philosophical engineering as well as between philosophical engineering and other kinds of engineerings.

true spirit of philosophical engineering, each of these projects *should* be subjected to further work of *reengineering*.

Moreover, there are several directions of further work that this dissertation opens. Below are some examples that I find particularly fruitful.

In connection to Chapter 1, one may want to further *substantiate the scope* of UA, especially as the field evolves. Or, one may want to identify and study particular *disputes* about UA.

In connection to Chapter 2, the lack of the literature that engages with the main research question at the same level of generality has, in part, resulted in presenting and discussing arguments in broad brushstrokes and offering only a blueprint of the method behind UA. That said, there is room for a more detailed argumentation as well as to consider and critically examine current and potential domain-specific approaches more carefully, especially as the field evolves. Furthermore, with respect to *pluralism about philosophical objects*, one may want to explore in more detail which of the existing as well as potential objects of philosophical engineering are tenable and compatible. With respect to the *five stage recursive model*, one may want to further fill out the details of the stages with respect to different criteria, such as methodological *efficiency* and *accuracy*. With respect to the claims about *engineering holism* and *division of engineering labour,* one may want to further explore the nature of *engineering* in order to inform us about the nature of the philosophic methodology as well as how to utilize different kinds of engineerings for the purposes of, for example, the strategic planning and the implementation stage of particular FOPs.

In connection to Chapter 3, one may want to engage in a large-scale implementation of the term 'philosophical engineering' for the general philosophical method behind UA. Furthermore, one may want to theorize further about the phenomena of *terminological dispute* and *terminological engineering*, as well as to engage in them.

In connection to Chapter 5, one may want to further defend a *Darwinian dilemma for realist theories of the nature of generics*, e.g. from the objections specific to generics as well as from the objections towards evolutionary debunking arguments. Furthermore, given the results of this chapter, one may want to investigate the prospects of: *antirealism* about genericity, *epistemicism* about the truth value of generics, *arbitrary reference* of generics, *semantic sovereignty* of generics.

In connection to Chapter 6, one may want to: develop other strands of *pragmatic engineering* of generics; look into the potential of engineering other propositional attitudes towards generics such as *desire* or *hope,* especially where changing *doxastic attitudes* such as *assertion* is not tenable; engage in further empirical work about, for example, the relation between generic judgments and generic propositions, and how to apply it to particular cases of engineering generics; or engage in FOPs of engineering *particular* generics.

Finally, one may want to research more closely further ramifications of: (i) philosophical engineering for *general philosophical methodology*; (ii) philosophical engineering of generics for *social reality*.

Part I

On the Nature of Philosophical Engineering: Towards Substantive and Terminological Dispute Resolution

We are not analyzing a world, we are building it.
We are not experimental philosophers, we are philosophical engineers.
(Berners-Lee 2003)

CHAPTER 1

THREE KINDS OF DISPUTES BEHIND THE UNITY ASSUMPTION

Abstract

In this chapter, I argue that those who theorize about the philosophical method behind FOPs submit to a common assumption about the methodology of FOPs. I take on board the following three assumptions that Cappelen and Plunkett (Cappelen and Plunkett 2020) claim are underlying FOPs (either disjunctively or in conjunction): (i) *Unifying Activity Assumption:* FOPs are instances of the same kind of activity; (ii) *Unifying Subject Matter Assumption:* FOPs share the same kind of subject matter; (iii) *Unifying Methodology Assumption:* FOPs share the same kind of philosophical methodology. I adopt a version of a *conjunction* of these assumptions which says that the *Unifying Methodology Assumption* consists in *Unifying Subject Matter Assumption* and *Unifying Activity Assumption*. I call this the 'Unity Assumption' (UA). Furthermore, I argue that despite a general background agreement about UA, there have been three kinds of disputes about UA: *merely verbal* (a disagreement concerning the *meaning* of the term(s) for the methodology behind UA), *substantive* (a disagreement concerning what the methodology behind UA *is*) and *terminological* (a disagreement concerning which *term* for the methodology behind UA *to use*).

1.1 Introduction

A growing group of philosophers has recently taken up an idea that a lot of philosophical endeavor *does not* (and, possibly, *should not*) happen merely at a level of *description*. As Burgess and Plunkett (2013a) point out: "claims about how one ought (or would do well) to think and talk are nearly as ubiquitous in philosophy as their descriptive counterparts, not to mention their prevalence in ordinary discourse" (Burgess and Plunkett 2013a, 1093). Instead, they argue that a lot of philosophical endeavor *does* (and, possibly, *should*) happen at a level of *evaluation* and *replacement* or *revision*. As Eklund (2014) stresses: "… while philosophers often have been concerned with our actual concepts or the properties or relations they stand for, philosophers should also be asking themselves whether these really are the best tools for understanding the relevant aspects of reality, and in many cases consider what preferable replacements might be" (Eklund 2014, 293). I will refer to an *engagement in this endeavor*[3] as a 'first order project' ('FOP').

In contemporary philosophical methodology, FOPs are often considered to be descending from Carnap's project of 'explication' which "consists in transforming a given more or less inexact concept into an exact one or, rather, in replacing the first by the second" (Carnap 1950, 3). However, we can see similar early tendencies in works of other philosophers such as e.g. Frege, Quine, Wittgenstein and Strawson to mention some. For example, Quine (1951) claims that in giving explicative definitions, "an activity to which philosophers are given, and scientists also in their more philosophical moments (…) the purpose is not merely to paraphrase the definiendum into outright synonym, but actually to improve upon the definiendum by refining or supplementing its meaning" (Quine 1951, 24–25). For a brief historical digression on the relevant work of other philosophers see Cappelen (2018, 24–27).

As an illustration of a FOP, consider Kevin Scharp's (2013) project of *replacing* the concept of TRUTH in the face of alethic paradoxes. After *evaluating* (with the help of an inconsistency theory that he develops) the (constitutive principles for) concept of TRUTH as inconsistent and, thus, unfit for our theoretical purposes (such as playing the central role in a semantic theory for natural language), Scharp proposes to *replace* the concept of TRUTH with the two concepts he designs especially for the purposes of avoiding the alethic paradoxes: ASCENDING TRUTH and DESCENDING TRUTH.

Below are some further examples[4] of what have recently been considered to be paradigmatic cases of engaging in FOPs, grouped[5] according to different contexts in which they may occur when applied:

---

[3] This endeavor has also been cashed out in terms of *conservative* versus *revisionary* conceptions of philosophical analysis (see Dutilh Novaes 2018), or conceptually *complacent* versus conceptual *skeptics* (see Cappelen 2018), or *creating* versus *discovering* (Nado 2019).

[4] Since space considerations preclude an exhaustive description of each of these examples, I will confine myself to a couple of footnote remarks about DSM CLASSIFICATION and PTSD.

[5] While grouping these cases, to a significant extent, I draw on Cappelen's (2018, Chapter 2), Ludlow's (2014, Chapter 2), and Chalmers' (2020) examples. However, I wish to stay neutral on which of these examples are actual cases of FOPs.

(i) *in philosophy*:

(a) to a particular concept such as: SENSE (Frege 1892); SUPERVENIENCE (Moore 1922; Hare 1952; Davidson 1970); TRUTH (Tarski 1933; Chihara 1979; Eklund 2002, 2014; Scharp 2007, 2013); INTENSION (Carnap 1947); GRUE and BLEEN (Goodman 1954); RIGID DESIGNATOR (Kripke 1971, 1980); IMPLICATURE (Grice 1975); KNOWLEDGE (Craig 1990); RACE (Appiah 1996); CONSCIOUSNESS (Block 1995); BELIEF (Clark and Chalmers 1998); CAUSATION (Woodward 2003); EPISTEMIC INJUSTICE (Fricker 2007); ALIEF (Gendler 2008); EXISTENCE (Thomasson 2015), KNOWLEDGE (Fassio and McKenna 2015); MISOGYNY (Manne 2017); CONCEPTUAL ENGINEERING (Isaac 2020), etc.

(b) more generally: explicating vague and indeterminate concepts (Carnap 1947, 1950); revisionary theories about moral language (Railton 1989, 1993; Joyce 2005); amelioration of concepts related to race and gender (Haslanger 1999, 2000, 2005, 2006, 2011, 2012); metalinguistic negotiations (Plunkett 2015); various accounts about how to fight effects of pernicious social kind generic statements (Haslanger 2011; Saul 2017; Leslie 2017), etc.

(ii) *in other domains of human inquiry:*

e.g. concepts such as MURDER, FETUS, IMMIGRANT, CHILD SOLDIER, PERSON, and TAX in law; DSM[6], PTSD[7], DEPRESSION in psychiatry; WEIGHT, KILOGRAM, METER, PLANET, SPACETIME in physics, etc.

---

[6] In psychiatry, DSM stands for 'American Psychiatric Association's *Diagnostic and Statistical Manual of Mental Disorders*.' DSM classification has gone through various revisions since its introduction, and, according to Van der Kolk (2014), it should have gone through more: "The foreword to the landmark 1980 DSM-III was appropriately modest and acknowledged that this diagnostic system was imprecise, so imprecise that it never should be used for forensic or insurance purposes. As we will see, that modesty was tragically short-lived." (Van der Kolk 2014, 42)

[7] Here is Van der Kolk's (2014) confession about how the lack of introducing definition of PTSD to DSM had profound negative effects:

"Each major diagnosis in the DSM had a workgroup responsible for suggesting revisions for the new edition. I presented the results of the field trial to our *DSM-IV* PTSD work group, and we voted nineteen to two to create a new trauma diagnosis for victims of interpersonal trauma: "Disorders of Extreme Stress, Not Otherwise Specified" (DESNOS), or "Complex PTSDâ" for short. We then eagerly anticipated the publication of the *DSM-IV* in May 1994. But much to our surprise the diagnosis that our work group had overwhelmingly approved did not appear in the final product. None of us had been consulted.

This was a tragic exclusion. It meant that large numbers of patients could not be accurately diagnosed and that clinicians and researchers would be unable to scientifically develop appropriate treatments for them. You cannot develop a treatment for a condition that does not exist. Not having a diagnosis now confronts therapists with a serious dilemma: How do we treat people who are coping with the fall-out of abuse, betrayal and abandonment when we are forced to diagnose them with depression, panic disorder, bipolar illness, or borderline personality, which do not really address what they are coping with?" (Van der Kolk 2014, 171)

And here is his illustration about how the introduction of a definition for PTSD to DSM had profound positive effects:

"The adoption of the PTSD diagnosis by the DSM III in 1980 led to extensive scientific studies and to the development of effective treatments, which turned out to be relevant not only to combat veterans but also to victims of a range of traumatic events, including rape, assault, and motor vehicle accidents. An example of the far-ranging power of having a specific diagnosis is the fact that between 2007 and 2010 the Department of Defense spent more than $2.7 billion for the treatment of and research on PTSD in combat veterans, while in fiscal year 2009 alone the Department of Veterans Affairs spent $24.5 million on in-house PTSD research." (Van der Kolk 2014, 189)

(iii) *in public debates:*

e.g. debates over concepts such as PERSON, MARRIAGE, RAPE, IMMIGRANT, REFUGEE, MONEY, TERRORIST, SANDWICH, ORGANIC, ATHLETE, etc., when seen as a form of *meaning negotiation* (Ludlow 2014).

(iv) *whenever we use language to communicate*:

e.g. as part of the phenomenon of *gradual semantic drift* (Dorr and Hawthorne 2014) and/or *contextual negotiation*[8] (Ludlow 2014), etc.

Some of these FOPs are result of a fairly recent developments whereas some of them have arguably been going on for decades, centuries, millennia even. However, only in recent years has there been an increasing interest among philosophers to *theorize* about FOPs. As a useful heuristic, one can distinguish between three different, nevertheless, interconnected levels of theorizing about FOPs: (i) *domain-specific theorizing about FOPs*; (ii) *general theorizing about FOPs*; (iii) *theorizing about the philosophical method behind FOPs*.

As an illustration of a *domain-specific theorizing* about FOPs one can, for example, consider Haslanger's (2012) theorizing about ameliorative projects about gender and race, or the above mentioned Scharp's (2013) theorizing which suggests the introduction of concepts of ASCENDING TRUTH and DESCENDING TRUTH. On the other hand, accounts of *general theorizing about FOPs* go by different names including 'conceptual engineering' (see Blackburn 1999; Brandom 2001; Scharp 2013; Eklund 2014, 2015; Cappelen 2018; Isaac 2020; Chalmers 2020), 'conceptual ethics' (see Burgess and Plunkett 2013a,b), 'verbal disputes' (see Chalmers 2011), 'ameliorative projects,' 'analytical projects' (see Haslanger 1999, 2000, 2006), 'revisionary projects' (see Railton 1989, 1993; Scharp 2007, 2013), 'explication' (see Carnap 1947; Quine 1951; Dutilh Novaes 2018; Brun 2016, 2017; Thomasson 2020), 'gradual semantic drift' (see Dorr and Hawthorne 2014), 'contextual and meaning negotiation' (see Ludlow 2014), and 'metalinguistic negotiations' (see Plunkett and Sundell 2013; Plunkett 2015). Most recently, attempts have also been made at *theorizing about the philosophical method behind FOPs* by recognizing that there is a general philosophical method behind FOPs worth theorizing about (see Cappelen 2018; Nado 2019; Isaac 2020; Burges, Cappelen and Plunkett 2020; Scharp 2013, 2020).

However, there is still a *foundational gap* about *theorizing about the philosophical method behind FOPs*. An excellent setup of foundational issues for conceptual engineering can be found in Issac (2020, see

[8] Ludlow's doctrine of Meaning Control says that "we (and our conversational partners) in principle have control over what our words mean. (…) If our conversational partners are willing to go with us, we can modulate word meanings as we see fit" (Ludlow 2014, 83).

esp. Section 2). For example, Isaac (2020, manuscript) detects and suggests an actionable plan for overcoming the foundational gap about the *subject matter* of this method[9] as follows:

> ... foundational gap consists in the *lack of any explicit account of the subject matter* of conceptual engineering (as already said, except Cappelen [2018], Machery [2017], and Sawyer [2018, 2020a,b]). Now, without such prior understanding of what it is about (i.e. without foundational theorization), *conceptual engineering will lack of any overall grip over its target domain and thus remain a useless, piecemeal method.* (…) *one should start theorizing its very subject matter*—and so overcome its foundational gap. (Isaac manuscript; italics mine)

In addition to the foundational gap about the *unifying subject matter* of the philosophical method behind FOPs, there is even bigger foundational gap about the *unifying activity* of the philosophical method behind FOPs. As Isaac (2020, 14) remarks: "Very few things have been said on the concept of 'engineering,' in the case of (conceptual) cognition, and for the purposes of conceptual engineering, except that it somehow involves an intentional and explicit design (e.g. Brun 2017: 2, 26 [see also Brun 2016: 1233 note 41])" (Isaac 2020, 14).

Part I of this dissertation contributes to *theorizing about the philosophical method behind FOPs*. In this chapter, I *elucidate the basic assumption* and *identify disputes* about the philosophical method behind FOPs. In Chapter 2, I work towards *clarifying the nature* and *offering a general framework* for the philosophical method behind FOPs. In Chapter 3, I show that the current terminology for the philosophical method behind FOPs is deficient and argue for the *introduction of a better terminology*.

In the rest of this chapter, I argue that those who theorize about the philosophical method behind FOPs submit to what I dub the 'Unity Assumption' (UA) (Section 1.2).[10] Finally, I argue that, despite a general background agreement about UA, we can distinguish between three kinds of disputes about UA: *merely verbal*, *substantive*, and *terminological* (Section 1.3).

1.2 The Unity Assumption

In recent years there has been an upsurge in recognizing that there is a research program behind FOPs. For example, Cappelen (2018, 9) claims that evaluative and revisionary activities with respect to our representational devices "constitute an interesting, unified category." However, he also points out that this is far from being obvious and that there is a reason for "why not a single book has been written that is focused exclusively on questions related to conceptual engineering: it is hard to see that there is a

[9] In his programmatic paper, after detecting this gap, Isaac (manuscript) argues for its relevance and offers a solution for how to fill this gap. In a nutshell, he theorizes about the subject matter of what he calls 'conceptual engineering' by "making a reflective use of conceptual engineering as a method for assessing and improving our representational apparatuses" (Isaac manuscript). His plan is to utilize the method of conceptual engineering in order to engineer the concept for the very same method, i.e. CONCEPTUAL ENGINEERING.

[10] For simplicity reasons, I will, henceforth, use the terms 'philosophical method' and 'methodological field' interchangeably given that the relevant authors use these terms both for the philosophical method and methodological field behind UA (cf. 'philosophical analysis' which stands both for the methodological field and philosophical method).

unified topic here. I think there is a unified topic—a rich field of research, even—but it takes work to clear the ground for it to become salient" (Cappelen 2018, 9–10). His book *Fixing Language: An Essay on Conceptual Engineering* (Cappelen 2018) is, arguably, one of the first systematic attempts to clear the ground for this kind of research field.

Another important attempt to clear the ground for this kind of research program is an edited volume on *Conceptual Engineering and Conceptual Ethics* (Burgess, Cappelen and Plunkett 2020). In its introduction, Cappelen and Plunkett (Cappelen and Plunkett 2020) offer a *top-down*[11] and *bottom-up characterizations* of what Cappelen prefers to call 'conceptual engineering' and Plunkett 'conceptual ethics.' The *top-down approach* and the *bottom-up approach* are two symmetric approaches that can be applied when theorizing about FOPs. Isaac (2020, 10) summarizes them as follows: "On one hand, top-down approaches, which require the theorization of what conceptual engineering is about (i.e. of its subject matter) prior to applying it, as a method, on specific study cases, and on the other hand, bottom-up approaches, according to which it is expected that "a theory of concepts [but what about 'engineering,' then?] will slowly emerge from a theory of conceptual engineering" (ConceptLab)" (Isaac 2020, 10).[12]

Cappelen and Plunkett (2020) offer two different characterizations of the field, suggesting that conceptual engineering and conceptual ethics can potentially be seen as two separated branches of philosophy. In particular, Cappelen (Cappelen and Plunkett 2020) draws on his previous work (see Cappelen 2018) and advocates for the field and the term 'conceptual engineering.' For the purposes of introduction, he allows for a broad scope[13] of the field that does not include only concepts but *representational devices* more generally: "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" (Cappelen and Plunkett 2020, 3). Cappelen's metaphilosophical considerations about conceptual engineering stem from his defense of conceptual engineering as more fundamental than other philosophical disciplines:

---

[11] See Burgess and Plunkett (2013a) for a similar approach.

[12] Even though Isaac (2020) sees them as compatible and complementary, Isaac argues for the primacy of the top-down approach "as opposed to the minimalism of bottom-up approaches to conceptual engineering, which advise remaining neutral with regard to the main foundational issues of the research program (e.g. Burgess and Plunkett 2013a: 1095; Plunkett 2015: 846)." (Isaac 2020, 10)

Moreover, Isaac (manuscript) believes that "only one of them, the so-called 'top-down' approach, may serve in order to purposefully address the main foundational issue in conceptual engineering (viz. that of its subject matter). Whereas its symmetric counterpart, the so-called 'bottom-up' approach, may well bring some foundational insights, but only in some kind of erratic and indirect way" (Isaac manuscript).

In this dissertation, I also give primacy to the top-down approach.

[13] The chief reason for that is that Cappelen beliefs that "since it's unclear and controversial what concepts are (and whether there are any), it's better to broaden the scope" (Cappelen and Plunkett 2020, 3).

> Conceptual Engineering is not usefully construed as a branch of any other part of philosophy. It will draw on insights from philosophy of language, philosophy of mind, epistemology, political philosophy, philosophy of science, ethics, and other field. That, of course, is also true about these other fields (i.e., they will draw on insights from each other). A case can, however, be made that Conceptual Engineering is prior to or more fundamental than all other philosophical disciplines. (Cappelen and Plunkett 2020, 4)

On the other hand, in the same introduction, Plunkett (Cappelen and Plunkett 2020), drawing on his previous work co-authored with Alexis Burgess (Burgess and Plunkett 2013a,b), offers the following summary of conceptual ethics:

> 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.) As the label suggests, some of the core issues in conceptual ethics concern *concepts* (assuming, for now, that there are such things). These include, centrally, normative issues about which concepts one should use (and why) and evaluative issues about which concepts are better than others (and why). (Cappelen and Plunkett 2020, 4)

Plunkett (Cappelen and Plunkett 2020) advocates for the idea that conceptual ethics is a separate branch of normative and evaluative inquiry, open to different philosophical commitments, and sees it as something that is worth both theorizing about as well as to engage in (as a method):

> Conceptual ethics is a branch of normative and evaluative inquiry, just as (at least certain parts of) epistemology, aesthetics, ethics, and political philosophy can be understood as branches of it. Thus, just as with other branches of normative and evaluative inquiry, people can approach conceptual ethics with very different philosophical commitments, very different views about how to make progress in it, and very different substantive views within it. Moreover, they can also approach it for very different reasons. For example, someone might be interested in conceptual ethics purely as an interesting part of philosophical theorizing. Or one might be interested in it because one is trying to actually change existing thought and talk. (Cappelen and Plunkett 2020, 5)

Most importantly, both Cappelen and Plunkett (Cappelen and Plunkett 2020) endorse the claim that there is *a generally shared assumption* that FOPs form a certain *unity*. As part of a bottom-up[14] way of introducing the topics of their volume, they claim that FOPs are about:

> … *that kind of activity or these kinds of issues*. The kind of thing they are doing, or the kinds of things they are discussing. The assumption then is that those activities form an interesting kind—a kind of *activity* or *method* or *subject matter* (or maybe all three at once). The idea of there being *some* sort of interesting kind here is a working hypothesis. One challenge for those working in the field is to try to substantiate it. Some chapters in this volume

[14] It is interesting to notice that even though they agree on the same bottom-up characterizations of the fields, they disagree on the top-down characterization. In particular, they see difference between conceptual engineering and conceptual ethics as not being merely terminological but also substantial since they each give a different top-down definition of the respective field. One could raise a question to what extent can they offer the same bottom-up characterization for something they take to be two different fields. However, I will not press this matters here further. Instead, I utilize their characterizations as two different versions of describing the methodological field behind FOPs. Also, it is worth noting that Burgess and Plunkett (2020) argue for a different solution for the relation between conceptual engineering and conceptual ethics.

support that assumption, and some argue against it. (Cappelen and Plunkett 2020, 19; italics mine)

One could, thus, argue that Cappelen and Plunkett (2020) bring out the following three assumptions underlying FOPs (either disjunctively or in conjunction): (i) *Unifying Activity Assumption:* FOPs are instances of the same kind of activity; (ii) *Unifying Subject Matter Assumption:* FOPs share the same kind of subject matter; (iii) *Unifying Methodology Assumption:* FOPs share the same kind of philosophical methodology. In this dissertation, drawing on Cappelen and Plunkett (Cappelen and Plunkett 2020), I adopt a version of a conjunction of these three assumptions which I call the Unity Assumption. The Unity Assumption says that the *Unifying Methodology Assumption* comprises *Unifying Subject Matter Assumption* and *Unifying Activity Assumption.*

In the final section of this chapter, I identify and bring to the foreground three kinds of disputes about the Unity Assumption.

1.3 Merely Verbal, Substantive and Terminological Disputes

1.3.1 Background

One of the prominent definitions of a *merely verbal dispute* offered by Jenkins (2014) characterizes a *merely verbal dispute* as follows: "(MVD+) A dispute is merely verbal iff: (i) the parties are engaged in a prima facie genuine dispute D on a certain subject matter S; (ii) the parties do not disagree on S; (iii) they appear to disagree on S because of divergent uses of language" (Jenkins 2014, 21).

Answering the question whether a certain philosophical dispute is merely verbal or substantive has become an important way of evaluating a philosophical debate.[15] As Belleri (2018, 691) points out: "Detractors of the debate may try to show that the participants in the discussion are merely "talking past each other." Their opponents interested in preserving the respectability of the controversy may try to show that the debate is either not merely verbal or, even if merely verbal, not problematically so" (Belleri 2018, 691).

In his influential paper on verbal disputes, Chalmers (2011) stresses the philosophical interest and role of verbal and substantive disputes for the philosophical discussion:

> First, they play a key role in philosophical method. Many philosophical disagreements are at least partly verbal, and almost every philosophical dispute has been diagnosed as verbal at some point. Here we can see the diagnosis of verbal disputes as a tool for philosophical progress. *If we can move beyond verbal disagreement to either substantive agreement or to clarified substantive disagreement, then we have made progress.* My own view is that the diagnosis of verbal disputes has the potential to serve as a sort of universal acid in philosophical discussion, either dissolving disagreements or boiling them down to the fundamental disagreements on which they turn. Second, verbal disputes are interesting as a subject matter for first order philosophy. Reflection on the existence and nature of verbal disputes can reveal something about the nature of concepts, language, and meaning.

[15] For more about merely verbal and substantive disputes see Hirsch 2005, 2008, 2009; Manley 2009; Chalmers 2011; Belleri 2018.

(Chalmers 2011, 517; italics mine)

Chalmers claims that "whether a dispute over S is verbal does not depend on S alone. It also depends on the parties to the dispute, and in particular on their background of agreement and disagreement." He also argues that: "The same sentence S can typically be the focus of both verbal and substantive disputes, depending on this background" (Chalmers 2011, 518). Furthermore, Chalmers (2011) endorses *conceptual pluralism* which he takes to lead to *semantic pluralism* and, ultimately, to *philosophical pluralism*. In particular, he claims that according to conceptual pluralism "there are *many* interesting concepts in the vicinity of philosophical terms such as 'semantic', 'justified', 'free', and very little depends on which one goes with the term" (Chalmers 2011, 539).[16] He further argues that the model of conceptual pluralism "leads to a sort of pluralism about the properties that these concepts pick out. For example, it naturally leads to semantic pluralism: there are many interesting sorts of quasi-semantic properties of expressions, playing different roles" (Chalmers 2011, 539–40). Lastly, he believes that this view also leads to a philosophical pluralism.[17]

Cappelen (2018, 23) expands Chalmers' idea about philosophical pluralism and argues that it has an important corollary aligned with the aims of the project of conceptual engineering: "the project of identifying all relevant concepts in the vicinity of a particular term and seeing to what use they can be put. This is a form of conceptual engineering" (Cappelen 2018, 23). More importantly, Cappelen (2018, 23) brings Chalmers' idea of verbal disputes in connection with philosophical pluralism, and argues that merely verbal disputes can often be seen as a way of a "communicative defect" that can be fixed by engaging in conceptual engineering: "Chalmers thinks that because we have not been aware of conceptual pluralism and haven't engaged in the kind of conceptual engineering that he advocates, very many philosophical debates have been pointless wastes of time. They have been exercises in what he calls 'verbal disputes'" (Cappelen 2018, 23).

Because of the amount of verbal disputes in philosophy Chalmers (2011) condemns most of philosophy as being 'pointless.' However, Cappelen (2018) argues that switching from a *pointless verbal dispute* to a *substantive* one via engaging in a certain form of conceptual engineering can be a way to improve this communicative defect. Cappelen (2018, 23) offers an example of a *verbal dispute* where two speakers use the same expression 'freedom' in order to refer to two different properties P and P*,

---

[16] However, he also admits that: "In cases where words have fixed connotations and associations, too, verbal issues often have serious practical import" (Chalmers 2011, 516–517). For more about the importance of our terminological choices see Chapter 3.

[17] "It leads to epistemic pluralism: there are many different epistemic relations, playing different roles. It leads to gene pluralism: there are many different things that deserve to be called 'genes', playing different roles. The same goes for confirmation pluralism, color pluralism, and so on. In fact, I am inclined to think that pluralism should be the default view for almost any philosophical concept. It may be that, as it happens, usage of a term such as 'gene' or 'confirmation' or whatever in our community is uniform enough that it has a single referent. But even so, there will be nearby possible communities, and probably numerous speakers within a community, who use the term in a different way, with equally interesting referents." (Chalmers 2011, 540)

respectively. If one of them says: 'Freedom is F,' and another one says: 'Freedom is not F' it seems like they are disagreeing. However, their disagreement dissolves once they realize that each of them uses 'freedom' to denote different properties.

In the rest of the Part I of this dissertation, I take on board Chalmers' (2011) notion of philosophical pluralism and its methodological corollary brought out by Cappelen (2018). In particular, by taking into consideration the role that different kinds of disputes as well as the background of agreement and the background of disagreement between different parties to the dispute play in philosophical discussion, I address two questions connected to a growing body of disputes about the philosophical method behind UA: (i) What kind of disputes are there about UA? (Section 1.3.2 and Section 1.3.3); (ii) How can we resolve these disputes? (Chapter 2 and Chapter 3).

1.3.2 Disputes Behind the Unity Assumption

In the vicinity of disputes about FOPs one can identify disputes both at the (*domain-specific and general)* theoretical and at the meta-theoretical level. It is also important to notice that there is an overlap in terminology that often stands both for meta-theorizing and (*domain-specific and general)* theorizing about FOPs (see Chapter 3, Section 3.2).[18] In Part I of this dissertation, I focus mainly on the prominent disputes at the meta-theoretical level or disputes about *the philosophical method behind FOPs* (see Section 1.1). In particular, I focus on disputes about the methodological assumption underlaying FOPs that I dubbed the 'Unity Assumption' ('UA') (see Section 1.2).[19]

In the rest of this section, by taking into consideration the role that disputing parties as well as the background facts about agreement and disagreement play in individuating disputes (see Chalmers 2011, 518), I spell out: (i) the background of *agreement* about UA, and (ii) the background of *disagreement* about UA, and distinguish between three kinds of disputes about UA: (iii) *merely verbal*, (iv) *substantive*, and (v) *terminological*.

(i) The Background of Agreement about UA: Parties to the dispute about UA *agree that there is* a unifying philosophical method behind FOPs that consists of a unifying subject matter and a unifying activity.

---

[18] This overlap has also been noticed by Isaac (2020, 4) who recognizes a multi-layered research program behind the term 'conceptual engineering' and draws a useful distinction between its reference to meta-theoretical vs. theoretical level: "'Conceptual engineering' is the name of a multi-layered research program. At a meta-theoretical level, it labels a renewed take on philosophy (more specifically, on philosophical methodology) conceived as the study of concepts. And at a theoretical level, it designates a fast-moving research field at the intersection of philosophies of language, mind, cognitive science, and epistemology, depending on the way in which concepts and cognition are more or less explicitly construed for the very purposes of conceptual engineering (see Sect. 3)." (Isaac 2020, 4)

[19] Recall, UA says that the *Unifying Methodological Assumption* consists in *Unifying Subject Matter Assumption* and *Unifying Activity Assumption*.

(ii) The Background of Disagreement about UA: Parties to the dispute about UA *disagree about* the unifying subject matter and the unifying activity of the philosophical method behind FOPs.

(iii) The Merely Verbal Dispute about UA: There is a disagreement among the parties to the dispute concerning the *meaning* of the term(s) for the philosophical method behind UA.

(iv) The Substantive Dispute about UA: There is a disagreement among the parties to the dispute concerning what the philosophical method behind behind UA *is*.

(v) The Terminological Dispute about UA: There is a disagreement among the parties to the dispute concerning which *term* for the philosophical method behind UA *to use*.

There is a sense in which disputes about the *meaning of the term* 'conceptual engineering' are merely *verbal*. For example, the meaning behind Scharp's (2013, 2020) and Isaac's (2020, manuscript)[20] use of term 'conceptual' in 'conceptual engineering' stands for *concepts* as opposed to Cappelen (see Cappelen 2018, 2020) who uses 'conceptual' to mean different entities including *representations* and/or *intensions* (for more details, see Chapter 3).

However, against a different background, namely by taking into consideration the above mentioned background of agreement and background of disagreement about UA, disputes about UA can be *substantive*. In particular, given the definition of UA (see Section 1.2) and its compositionality, we can distinguish between two underlying substantive disputes about UA:

(i) The Unifying Subject Matter Dispute about UA: There is a substantive disagreement among the parties to the dispute about what is the unifying subject matter of the methodology of FOPs. For example, consider a dispute whether the unifying subject matter is *concept psychological construed* (see Isaac 2020, manuscript[21]), or *concept philosophically construed* (see Scharp 2013, 2020[22]), or *intensions* (see Cappelen 2018), or *beliefs* (see Greenough 2020), etc.

---

[20] There is, perhaps, a room for detecting a further dispute between Scharp (2013, 2020) and Isaac (2020) given the particular meanings for concepts they use.

[21] Isaac establishes that "conceptual engineering should be about concepts, psychologically theorized, as multiply realizable functional kinds" and that he has "secured and justified the broadest scope, the strongest impact, and the highest flexibility for conceptual engineering as a method for the cognitive optimization of ours representational apparatuses" (Isaac manuscript).

[22] "I have laid out the radical therapeutic vision of what philosophy is all about. I have presented conceptual engineering as one aspect of a proper philosophical methodology. We should take an active role in altering and improving our conceptual scheme. I have also advocated a scientific element in this philosophical methodology that I have called metrological naturalism." (Scharp 2020, 409)

(ii) The Unifying Activity Dispute about UA: There is a substantive disagreement among the parties to the dispute about what is the unifying activity of the methodology of FOPs. For example, consider a dispute whether the unifying activity is *change in constitutive principles* (see Scharp 2013), *reference change* (see Cappelen 2018), *engineering* (see Chalmers 2020; Isaac manuscript; Greenough manuscript; Fuš Chapter 2, Section 2.3), etc.

Lastly, dispute about UA can be *terminological*. In Chapter 3, I introduce a *terminological dispute* as a novel kind of dispute that has not yet received appropriate attention. In a nutshell, I define a *terminological dispute* as follows: A dispute over a term T is terminological when, for some T, the parties *do not disagree* about the meaning behind T, rather the dispute over T arises wholly in virtue of the disagreement about whether to use T rather than an alternative term, A. For example, consider a dispute about which of the terms (if any), 'conceptual engineering' or 'conceptual ethics,' is a more adequate terminological choice for the philosophical method behind UA.

1.3.3 Towards Resolving Disputes

Merely verbal disputes may be an indicator that the participants of the discussion are talking past each other. In fact, diagnosing a certain dispute as a merely verbal has often been seen as an "impediment to understanding" (Chalmers 2011, 517) and an obstacle to move towards a substantial dispute (see also Hirsh 2009; Thomasson 2016).

The first step of a *merely verbal dispute resolution* should be for all the parties to the dispute to exercise the *principle of charity* (see Hirsch 2005, 2008, 2009; Belleri 2018), after which they should come to a conclusion that they were all speaking the truth in their own versions of their own language or "idiolect." The next step should then be for them to settle on the meaning of the disputed term. Similarly, in the case of resolution of the merely verbal dispute about UA, as a first step, the parties to the dispute should realize that their disagreement was merely verbal because they used the same term whereas they ascribed different meanings to it. The next step of the merely verbal dispute resolution about UA should be to settle on which meaning they should ascribe to the term about UA.

However, resolving a merely verbal dispute by settling on one of the proposed meanings for the term may not automatically settle the *substantive* nor *terminological dispute*.[23] In the case of disputes about UA, one can argue that even if the parties to the dispute settled a merely verbal dispute, i.e. even if they, for example, settled on the *meaning* of the terms they use, both substantive and terminological

[23] Here is an example that Chalmers (2011, 518) offers in order to illustrate how a dispute can be both verbal and substantive depending on the background: "… against the background of agreement on the motion of a squirrel, a dispute over 'The man goes around the squirrel' may be verbal, but against a different background (perhaps we agree that the man goes around some animal but disagree about whether it is a squirrel or a mouse), the dispute may be substantive. So we cannot just speak of a "verbal question," independent of context." (Chalmers 2011, 518)

disputes about UA would still remain unsettled. In particular, in the case of a *substantive dispute about UA*, the parties may still disagree what the philosophical method behind UA in fact *is*. And, in the case of a *terminological dispute about UA*, the parties may still disagree on what *term* is best to *use* for the philosophical method behind UA.

In Chapter 2 and Chapter 3, I offer resolutions for the *substantive* and the *terminological* dispute behind UA, respectively.

CHAPTER 2

SUBSTANTIVE DISPUTE RESOLUTION OF THE UNITY ASSUMPTION

Abstract

In this chapter, I offer a resolution of the *substantive* dispute behind UA. I argue that none of the current proposals can settle the substantive dispute about UA. The reason for that is, I argue, that the so-far proposed theories for the philosophical method behind UA cover *the unifying subject matter* and *the unifying activity scope* only partially. This further suggests that a substantive disagreement could not be resolved by settling on one of the current proposals. Instead, by arguing for a more plausible scope of the unifying subject matter and the unifying activity behind UA, I offer another way to resolve the substantive disputes over the unifying subject matter and unifying activity that comprise the unifying methodology of FOPs. I engage in a *substantive dispute resolution* of UA as follows. First, I argue that we should consider *being philosophical* as the unifying subject matter behind UA. More specifically, I suggest that what makes the method behind UA *philosophical* is that philosophy as its *subject matter* is being engineered. This suggestion relies on the assumption that philosophy as a FOPs' subject matter can be engineered via engineering philosophical objects. Second, I argue that we should consider *engineering* as the unifying activity behind UA. More specifically, I suggest that what makes the method behind UA *engineering* is that it comprises a *five-stage* (*identification, evaluation, strategic planning, implementation, re-evaluation*) *recursive activity*. This suggestion takes engineering activity in a non-metaphorical sense and endorses engineering as a philosophical method that bears a *family resemblance* and is in *continuity with other engineering fields*. In a nutshell, *philosophical engineering* is engaging in a five-stage recursive activity that can operate on different philosophical objects.

2.1 Preliminaries

Throughout the twentieth century, philosophers belonging to the various schools of philosophy argued for philosophical analysis as being the core part of analytic philosophy. The broad dilemma for philosophical analysis has been whether philosophy is aiming at understanding linguistic expressions and concepts on the one hand, or truths and facts about the external world on the other. For example, Cappelen (2018, 47) contrasts two competing descriptivist camps about the nature of philosophy: one aiming at describing concepts (e.g. Dummett 1978; Jackson 1998) and another aiming at describing extensions of those concepts (e.g. Williamson 2007; Kornblith 2002). As a result of this dilemma, different conceptions of philosophical analysis that aim at revealing the relationship between language, logic, thought and reality have been developed.[24] Nevertheless, as Beaney (2014) notices: "Analytic philosophy is alive and well precisely because of the range of conceptions of analysis that it involves. It may have fragmented into various interlocking subtraditions, but those subtraditions are held together by both their shared history and their methodological interconnections" (Beaney 2018). Furthermore, in his paper 'What Is a Philosophical Analysis?,' King (1998) raises substantive questions about the nature of philosophical analysis:

> What is it that is being analysed (i.e. what sorts of things are the objects of analysis)? What sort of thing is the analysis itself (a proposition? sentence?)? Under what conditions is an analysis correct? How can a correct analysis be informative? How, if at all, does the production of philosophical analyses differ from what scientists do? (King 1998, 155)

A working hypothesis of this chapter is that *substantive questions* about the nature of *philosophical analysis* translate to analogous substantive questions about the nature of a philosophical method behind UA. As a precursor for answering substantive questions behind UA, I consider Cappelen's (2020) methodological account that he dubs the 'Master Argument for Conceptual Engineering.'[25] Cappelen (2020) does not claim the originality of the argument. Instead, he wants to articulate the most general (not domain-specific) version of the argument (Cappelen 2020, 134):

> As a heuristic it's useful to think of conceptual engineering as having two parts: the general theory and the specific applications. We should expect a two-way interaction: the general theory will inform the specific cases

[24] E.g. Frege's and Russell's notions of logical analysis; Moore's semantic (and of the world) analysis; Carnap's quasi-analysis and explication; Wittgenstein's analysis of the ultimate constituents of propositions and the nature of the world itself; Oxfords's linguistic and connective analysis; logical and metaphysical analysis of the Cambridge School of Analysis (see Beaney 2014, Soames 2003). For a quick overview, see Beaney's (2018) SEP entry on *Conceptions of Analysis in Analytic Philosophy*.

[25] "(1) If W is a word that has a meaning M, then there are many similar meanings, M1,M2,...,Mn, 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, she should assess and ameliorate the meanings of central terms." (Cappelen 2020, 134)

> and the specific cases will inform the general theory. It should also be clear from the discussion above that *there can be many frameworks for thinking about conceptual engineering*. What one takes conceptual engineering to be (when thinking about both the general theory and specific cases) will be shaped in large part by what one takes meanings and concepts to be, what one assumes about metasemantics, and what one takes to be conceptual defects and virtues. (Cappelen 2020, 150; italics mine)

Motivated by Cappelen's (2020) Master Argument, I adapt two general principles in offering a resolution to a substantive dispute about UA: (i) *The Plurality Principle*: takes on different conceptions of philosophical method behind UA. In result, this will contribute to the *scope broadness* for the proposed philosophical method behind UA; and (ii) *The Unity Principle*: accommodates the commonalities among different conceptions of philosophical method behind UA. In result, this will contribute to the *scope restriction* for the proposed philosophical method behind UA.

Lastly, I take on board three core assumptions presented in Chapter 1: (i) I adopt the view that FOPs (First Order Projects, see Chapter 1) share a philosophical method, formulated as the Unity Assumption (UA); (ii) I take current theorizing about UA as data points; (iii) I rely on the background of agreement among the parties to the substantive dispute about UA. With the above preliminaries in place, in the rest of this chapter, I work towards resolving the *unifying subject matter dispute* (Section 2.2) and the *unifying activity dispute* about UA (Section 2.3).

2.2 Resolving the Unifying Subject Matter Dispute

In this section, after considering some of the current proposals (Section 2.2.1) and suggesting the most plausible unifying subject matter scope for FOPs (Section 2.2.2), I suggest how to resolve the unifying subject matter substantive dispute about UA (Section 2.2.3).

2.2.1 Current Proposals

In recent years, substantive dispute about *the subject matter* behind UA has been in the main focus of current debates about UA. There has been an upsurge in different *conceptions* of what the unifying subject matter behind UA is. The unifying subject matter is often articulated as different 'entities' or 'categories' that are considered to be 'targets'[26] or 'objects' of FOPs, hereinafter referred to as *FOP objects*. Below I offer two recent summaries of the state of the art of this fast-moving research field.

First. In *Fixing Language*, Cappelen (2018) considers a possibility of different FOP objects: "What is Conceptual Engineering about? Does it concern *concepts, meanings, extensions, intensions, or something else*?" (Cappelen 2018, 50; italics mine).[27] Furthermore, he raises questions about their nature: "What are

[26] See Nado 2020.

[27] To compare, philosophers working in the area of standard Philosophical Analysis have questioned the traditional idea that concepts are the (only) objects of standard philosophical analysis and have offered answers to the standard Philosophical Analysis without invoking concepts. For example, for King (1998) properties and relations play this role.

the things we criticise and improve? So far I've called them concepts, but it is not clear that that is the right label for the targets of engineering.[28] If we engineer concepts, what are they, and how do we go about changing them?" (Cappelen 2018, 51). Cappelen admits that there "are a plethora of alternative approaches to conceptual engineering. Many of these are embedded in different theories about language, concepts, semantics, pragmatics, and communication" (Cappelen 2018, 161). However, he also admits that we are still far from clear "on the exact nature of the entities being engineered. (…) The first item on the agenda for such views [viz. the ones that take conceptual engineering to be about concepts] should be to specify what concepts are, and then present an account of how concepts so construed can be engineered" (Cappelen 2018, 141).

Second. More recently, in the Introduction to *Conceptual Engineering and Conceptual Ethic*s, Cappelen and Plunkett (2020) spell out some possible candidates for 'representational devices,' a more general term they deliberately use for the FOP object: "What are the relevant representational devices? Possible answers include: concepts (as they are construed in some part of psychology or philosophy), lexical items, and the semantic values of lexical items. A closely connected cluster of questions concerns whether they are in language or thought or both" (Cappelen and Plunkett 2020, 3). More importantly, they admit that: "Different conceptual engineers will give different answers and that will have enormous implications for how the field is understood and practiced" (Burgess, Cappelen, Plunkett 2020, 3).

As part of a search for the most plausible subject matter candidate, I offer a brief attempt at a tentative taxonomy of some of the current *conceptions* of FOP objects. I divide them in three broad categories:[29]

(i) Concepts:[30] FOP objects are concepts construed as philosophical and/or psychological entities.[31]

---

[28] For example, Cappelen's own view, the Austerity Framework, "understands conceptual engineering as a form of reference change" (Cappelen 2018, 51).

[29] Cf. Cappelen identifies three broad kinds of conceptual engineers, namely *topic-improving* engineers, *semantic* engineers and *lexical effect-exploiting* engineers (aka Exploiters), which are not (or at least, in certain cases, not directly) about concepts (see Cappelen 2018). On the other hand, Isaac (2020, manuscript) roughly distinguishes between three broad categories that I partly adopt below: *philosophical* or *semantic approach to concepts*, *psychological* or *cognitive approach to concepts*, and *unprincipled* views.

[30] Here I draw on Isaac (2020) about "a widely acknowledged, and yet controversial, distinction between philosophical and psychological approaches to concepts (Löhr 2018; Machery 2009; Peacocke 1992; Rey 1985, 2010)" (Isaac 2020, 11).

[31] E.g. Brun (2016, 1217) takes concepts to be word/rules-of-use pairs, Prinzing (2018, 858) takes them to be cognitive tools, and Kitsik (2018, §3) takes concepts to be ways of using expressions in thought or talk.

(a) Philosophical or semantic approach to concepts;[32]

(b) Psychological or cognitive approach to concepts;[33]

(c) Dual view of concepts.[34]

(ii) Unprincipled views:[35] it does not matter whether FOP objects are concepts or any other kind of representational devices.[36]

(a) *Linguistic elements:* e.g. linguistic meaning, terms, lexical items, etc.

(b) (*Meta)semantic elements:* e.g. extensions, reference, truth, intensions (through e.g. metasemantics[37]), etc.

(c) *Pragmatic elements:* e.g. speaker-meaning,[38] propositional attitudes, speech acts,[39] etc.

---

[32] "Typically, philosophical approaches to concepts take them to be semantic entities endowed with 'semantic structure' (Margolis and Laurence 2010, 219), and focus on their semantic constituency in order to deliver (a priori, analytic) truths about the world (e.g. Jackson 1994, 1998; Peacocke 1992; Pitt 1999)" (Isaac 2020, 11). Cf. Brun 2016; Burgess and Plunkett 2013a; Chalmers 2011; Plunkett 2016; Sawyer 2018; Scharp (2013, 2020); Simion 2018; Thomasson 2020.

[33] "By contrast, psychological approaches typically take concepts to be cognitive entities endowed with 'processing structures' (Margolis and Laurence 2010, 219), and focus on the role they play in our higher cognition in order to deliver empirical propositions about the mind (e.g. Murphy 2004)" (Isaac 2020, 11). Cf. Isaac (2020, manuscript), Machery (2017).

[34] E.g. according to Isaac (2020, 11), Plunkett (2015) and Prinzing (2018) could be seen as proponents of concepts having both semantic and processing structures. Similarly, Koch (2020a) proposes a dual content view by understanding concepts as having two (interrelated) kinds of contents: referential content and cognitive content.

[35] E.g. Ball 2020; Burgess and Plunkett 2013a; McPherson and Plunkett 2020; Nado 2019; Prinzing 2018; Shields 2020; Sterken 2020; Tanswell 2017.

[36] "I have, for example, variously referred to the 'entities', 'concepts', 'terms', 'topics', and 'subject matters' speakers are attempting to get right. How speakers will choose to characterize the precise nature of the "something" they are trying to get right will depend on the specifics of each case and is not something we should try and dictate from the armchair." (Shields 2020, 14)

[37] See Cappelen 2018.

[38] See Pinder 2019.

[39] See Chapter 6 of this dissertation.

(d) Epistemic elements: e.g. revelation,[40] etc.

(e) Miscellaneous philosophical elements: e.g. theses,[41] topics,[42] roles,[43] properties,[44] theories,[45] functions,[46] beliefs,[47] ideas,[48] conceptions,[49] classification procedures,[50] etc.

(iii) Derivative views: not *directly* FOP object candidates but only *indirectly* via utilizing some of the above mentioned FOP object candidates.

---

[40] E.g. Haslanger (2006) takes ameliorative analysis to be a process of *uncovering* which concepts we have been employing all along, rather than as one involving conceptual *revision*. As she puts it: "one potential advantage of a constructionist account is that it does not simply deny the existence of the allegedly natural category and substitute another (possibly social category) in its place, but it also—at least in the best cases—provides a diagnosis of our role in bringing about the effects that appear to us (mistakenly) as natural, together with an explanation of the illusion. In such cases, the self-deception involved when we mean something, and yet mask that meaning to ourselves, is laid bare. Such unmasking can be an important step in motivating social change" (Haslanger 2006, 94).

Cf. also Sterelny (2012) who defined 'epistemic engineering' as "organizing our physical environment in ways that enhance our information-processing capacities about some target domain" (Sterelny 2012, xii).

[41] "Theses have a certain primacy in philosophy, or at least in the sort of philosophy aimed at discovering truths. In this sort of philosophy, conceptual engineering is largely driven by thesis engineering." (Chalmers 2020, 13).

[42] See Cappelen 2018.

[43] "Where practical projects are concerned, normative and evaluative theses will often be central to the philosophical grounding of these projects. But perhaps aiming for true theses doesn't capture all that's going on. So perhaps more generally, one should speak of the primacy of roles rather than the primacy of theses, and understand the fruitfulness of a concept as tied to the roles it can play." (Chalmers 2020, 14)

[44] A view that Scharp (2020) does not endorse but rather anticipates: "Philosophy isn't the study of concepts at all, so it cannot be the study of what have turned out to be inconsistent concepts. Philosophers do on occasion study concepts, but only as one item among many in other things in the world. For example, there is a difference between the concept of truth and truth itself. Truth is, presumably, a property that things like sentences or theories or propositions can have, whereas the concept of truth is something like a mental representation or a constituent of thought or some other kind of thing that people grasp or possess or understand. Philosophy isn't the study of *the concept* of truth or *the concept* of knowledge or any of the other concepts. Instead, philosophy is the study of certain *phenomena*, like truth, knowledge, freedom, justice, and the rest" (Scharp 2020, 413).

[45] See Sawyer 2018.

[46] As Nado (2020) points out: "Functionalist accounts vary on just what a concept's function is. Some accounts tie a concept's function to its semantic features (e.g. Brigandt (2010) and Fisher (2015)). Thomasson (2020) suggests making use of Millikan's notion of 'proper function', while Haslanger (2020) plumps for a version of the 'systems' approach to functions" (Nado 2020, 11). For Nado (2019, 2020) these are practical roles (see Nado 2020).

[47] See Pollock 2020.

[48] See Greenough (2020): "… we merely need to revise or replace our ideas, beliefs, theories, and conceptions about the things picked out by those words. Idea Engineering is all we really need" (Greenough 2020, 207).

[49] "Although a conception may sound very similar to a theory, it need not be identified with it; we may think of theories as better-developed, better-structured and better evidentially supported versions of conceptions, and of conceptions as generally less polished, less tested and less "mature" versions of theories." (see Belleri manuscript)

[50] See Nado 2020.

(a) Effects:[51] e.g. social, moral, political,[52] cognitive, lexical,[53] etc.

(b) Reality:[54] e.g. social facts,[55] behavior, etc.

A few remarks about this tentative taxonomy are in order. First, it is worth stressing that since we are in early stages of theorizing about the general method behind UA, building a taxonomy of FOP object candidates can be seen as an ongoing project, both with respect to which FOP object candidates to include as well as how to categorize them. What I offered above is far from ideal and is intended to serve merely as a tentative guide. Furthermore, giving a complete taxonomy goes beyond the scope of this chapter. Instead, for the purposes of this chapter, it will be enough to assume that there are different FOP objects candidates. Finally, it has been, especially within philosophical analysis, long taken for granted that different conceptions of FOP objects *compete* with each other. More recently, however, while arguing for a certain kind of FOP object as being the primary or the most significant, many increasingly (at least implicitly) endorse the possibility of other FOP objects. In the rest of this chapter, I further elucidate and develop the latter strand, i.e. I advocate for the Pluralist Approach that allows for different conceptions of FOP objects, none of which is independently privileged.[56]

---

[51] See Cappelen 2018.

[52] For social, moral and political effects in particular, see Chapter 6 of this dissertation.

[53] For cognitive and lexical effects in particular, see Chapter 3 of this dissertation.

[54] "the idea that philosophy consists in [the] investigation of our concepts is [...] deeply mistaken. [...] [T]he subject matter of ethics is the right and the good, not our concepts of them. The subject matter of philosophy of mind is the mind itself, not our concept of it. And the subject matter of epistemology is knowledge itself, not our concept of knowledge." (Kornblith 2002, 1)

"The primary concern of epistemology is with the nature of knowledge, not with the nature of the concept of knowledge. If knowledge were in fact identical with justified true belief, that would be what mattered epistemologically, irrespective of the conceptual possibility of their non-identity." (Williamson 2007, 206)

[55] See Cappelen 2018.

[56] Here is a quick parallel with Haslanger's (2006) view that different kinds of philosophical analysis she distinguishes are closely connected and are co-workers rather than competitors. In particular, Haslanger argues that: "Conceptual, descriptive and ameliorative projects cannot, of course, be kept entirely distinct, but they have different subject matters and different goals" (Haslanger 2006, 96).
Furthermore, she claims that "which approach is better will depend on the case at issue, and the betterness will depend on semantic, pragmatic and political considerations" (Haslanger 2006, 94). However, she claims that philosophers have sometimes been engaged in different kinds of projects of philosophical analysis with respect to the same concept: "Some authors are engaged in a conceptual project, attempting to explicate our ordinary understanding of race (Appiah 1996, Zack 1997, Hardimon 2003, Mallon 2004); others are attempting to determine what, if any, natural kind we are referring to by our racial terms Appiah 1996, Kitcher 1999, Andreason 2000, Zack 2003); others have pursued genealogy (Omi and Winant 1994); still others are invested in what I call ameliorative projects, raising normative questions about how we should understand race, not only how we currently do (Gooding-Williams 1998, Alcoff 2000)"(Haslanger 2006, 96). Finally, she believes that conceptual and ameliorative projects can come together when: "(a) the concept we take ourselves to be employing, (b) the concept that best captures the type we are concerned with, and (c) the type we ought to be concerned with coincide. In such cases the conceptual and ameliorative projects yield the same concept. It is a mistake, then, for those engaged in conceptual analysis to dismiss other forms of analysis, with the thought that only the conceptual project can discover 'our' concept. (see Mallon 2004, Hardimon 2003)" (Haslanger 2006, 96–97).

2.2.2 The Most Plausible Unifying Subject Matter Scope: The Pluralist Approach to FOP Objects

In this section, I introduce and argue for the *Pluralist Approach to FOP Objects* as the most plausible *unifying subject matter scope* behind UA. The proposed Pluralist Approach to FOP Objects comprises the *widest subject matter scope* behind UA. As a restriction on the subject matter scope, i.e. on what can fall under the *FOP object*, I utilize the *bottom-up (descriptive) approach,* by taking into account the commonalities among the current theorizing about *objects* of the unifying subject matter behind UA as data points (see Section 2.2.1). Furthermore, I also utilize the *top-down (normative) approach*, namely I spell out what the *object* of the unifying subject matter behind UA *should be*, by introducing and supplementing certain *objects* that have not (yet explicitly) emerged from the bottom-up approach. In Section 2.2.3, I then suggest for *being philosophical* as the unifying subject matter *category*.

2.2.2.1 Scope Candidates

Consider the following three ways of construing the subject matter scope behind UA.

Subject Matter Scope (i): If FOPs are conceived in the way that they can include *only one of the currently proposed FOP objects*, then the subject matter dispute about UA would be resolved by settling on *one* of the currently proposed FOP objects as its subject matter. In practice, this would mean that the subject matter behind UA is, for example, either a concept, or an intension, or a term, or a topic, or a belief, etc. In result, all current and potential projects with this particular object (and the same unifying activity behind UA, see Section 2.3) would count as FOPs.

Subject Matter Scope (ii): If FOPs are conceived in the way that they can include *more than one of the currently proposed FOP objects*, then the subject matter dispute about UA would be resolved by settling on *more than one of the currently* proposed FOP objects as its subject matter. In practice, this would mean that, for example, a concept, an intension, a term, a topic, a belief, etc., would jointly be objects that fall under the subject matter behind UA. In result, all current and potential projects with these different objects which share the same subject matter (and the same unifying activity behind UA, see Section 2.3) would count as FOPs.

Subject Matter Scope (iii): If FOPs are conceived in the way that they can include *more than one FOP object which can include both those objects that have currently been proposed as FOP objects as well as those objects that have not (yet) been proposed as FOP objects*, then the subject matter dispute about UA can be resolved by settling on *more than one FOP object, including both currently proposed objects as well as*

objects that can be proposed as being the subject mater about UA. In practice, this would mean that, for example, not only a concept, an intension, a term, a topic, a belief, etc., would jointly be objects that fall under the subject matter behind UA but also objects that have not yet been considered as FOP objects e.g. propositions, propositional attitudes, judgments,[57] sentences, speech acts, etc., would also fall under the subject matter behind UA (for restriction of this scope see Chapter 4, Section 4.1.1). In result, all current and potential projects with these different objects that share the same subject matter (and the same unifying activity behind UA) would count as FOPs.

2.2.2.2 The Widest Scope Rationale

For the purposes of evaluating the above mentioned subject matter scope candidates, I utilize the *Widest Scope Rationale* and I argue that the *Subject Matter Scope (iii) is the most plausible* candidate.

The Widest Scope Rationale. According to UA, the unifying subject matter and the unifying activity are constitutive of FOPs belonging to the same methodological family. (Analogously, in the case of conceptual analysis, the unifying subject matter object is a concept and the unifying activity is analysis. Analyzing concepts is what makes a certain project to belong to the methodological family of conceptual analysis.) Belonging to the same methodology can have various important impacts for FOPs. For instance, being treated as part of the same methodology can contribute to the successfulness of implementation of FOPs. Successful implementation is important because FOPs are significant.[58] Given that, the widest unifying subject matter scope (and the widest unifying activity scope[59]) for FOPs should include the greatest amount of FOPs that can fall under the same methodology and, thus, have the benefits of the same methodology.

However, in order to avoid the danger of the *Widest Scope Rationale* to be employed too broadly (i.e. to mis-categorize some projects as FOPs and apply methodology inappropriately), for the present moment, one should assume that all the three of the above mentioned scopes fall under to the same *category* that I defend as the unifying subject matter in Section 2.2.3. With this in mind, next, I evaluate the above mentioned subject matter scope candidates with respect to how the resolution of a substantial disagreement about the subject matter of UA impacts the FOPs' scope size.

The *Subject Matter Scope (i)* allows only for one FOP object. Call this *singularism about FOP objects*. In

---

[57] See Chapter 6.

[58] E.g. Cappelen (see 2018, 39–46) has stressed the significance of conceptual engineering in its impact on the world by offering three arguments: two Ontological and the Prudential Argument. Very briefly, the two *Ontological Arguments* state that: (i) language plays the constitutive role in social reality (building on Searle 1995), and also, more radically, (ii) conceptual amelioration is an amelioration of the world (building on Haslanger 1999, 2000). *The Prudential Argument* rests on the Revisionist's Basic Assumption and stresses the importance of assessing and improving of our representational devices (see Cappelen 2018, 40). Furthermore, Burgess and Plunkett (2013a) believe that since "cognitive errors often have negative practical consequences, one can start to see why we might be inclined toward certain normative views about conceptual or semantic matters" (Burgess and Plunkett 2013a, 1093).

[59] I apply the same rationale in the case of the most plausible scope of the unifying activity behind UA (see Section 2.3).

particular, *Subject Matter Scope (i)* resolves the substantial disagreement about the subject matter behind UA by narrowing down the subject matter scope for FOPs to only one object, e.g. concepts. Under the above assumption that all the three scope candidates fall under to the same category (which I defend as the unifying subject matter in Section 2.2.3), *Subject Matter Scope (i)* unnecessarily reduces the number of projects that can count as FOPs, and is, thus, *too narrow* subject matter scope for FOPs.

The *Subject Matter Scope (ii)* and the *Subject Matter Scope (iii)* allow for multiple FOP objects. Call this *pluralism about FOP objects*. Both of the scope candidates resolve the substantial disagreement about the subject matter behind UA by extending the number of FOP objects.

The *Subject Matter Scope (ii)* takes (at least some of the) currently competing FOP objects as candidates for belonging to the same methodological family (as long as they share the same unifying subject matter and activity). In result, according to the *Subject Matter Scope (ii)*, projects that have different FOP objects and share the same unifying subject matter (and the same unifying activity) share the same methodology. Consequently, under the above assumption that all the three scope candidates fall under to the same category (which I defend as the unifying subject matter in Section 2.2.3), *Subject Matter Scope (ii)* allows for a wider subject matter scope for FOPs than *Subject Matter Scope (i)*. This, given the Widest Scope Rationale, makes the *Subject Matter Scope (ii)* a *more plausible* subject matter scope candidate than the *Subject Matter Scope (i)*.

The *Subject Matter Scope (iii)* takes as FOP objects candidates not only (at least some of) those objects that have already been proposed but also considers objects that are possible but have not been proposed yet as candidates for belonging to the same methodological family (as long as they share the same unifying subject matter and activity). This means that both (at least some of the) current as well as (at least some of the) possible projects which include objects that share the same unifying subject matter (and the same unifying activity) but have not yet been proposed, can count as FOPs which share the same methodology. *Subject Matter Scope (iii),* thus, includes wider subject matter scope than the previous two. Consequently, under the above assumption that all the three scope candidates fall under the same category (which I defend as the unifying subject matter in Section 2.2.3), *Subject Matter Scope (iii)* resolves the subject matter dispute behind UA in a way that gives us the widest scope impact for FOPs. This, given the Widest Scope Rationale, makes the *Subject Matter Scope (iii) the most plausible* subject matter scope candidate for the methodology behind UA.

2.2.2.3 Scope Candidate Impact

Here is a brief illustration of the *scope impact* of the above mentioned *subject matter scope candidates*. For simplicity, I will narrow down *current* dispute about the subject matter scope to two *current* FOP object candidates: *intensions* (see Cappelen 2018) and *concepts* (see Scharp 2013, 2020), and consider one *potential* FOP object candidates: *speech acts* (see Chapter 6).

If we adopted the *Subject Matter Scope (i)*, the substantive dispute resolution would be settled by taking *either* concepts *or* intensions as FOP objects. In other words, *either* projects that take intensions as objects *or* projects that take concepts as objects (and also share the same unifying activity) would count as FOPs.

On the other hand, if we adopted *Subject Matter Scope (ii),* the substantive dispute resolution would be settled by taking *both* concepts *and* intensions as FOP objects. In other words, *both* projects that take intensions as objects *and* projects that take concepts as objects (and also share the same unifying activity) would count as FOPs.

Lastly, if we adopted *Subject Matter Scope (iii),* the substantive dispute resolution would be settled by taking *not only* concepts and intensions, *but also* speech acts to count as FOPs. In other words, projects that take *intensions* as objects, projects that take *concepts* as objects, and projects that take *speech acts* as objects (and also share the same unifying activity) would all count as FOPs. Most importantly, in addition to being able to successfully account for more projects as FOPs, *Subject Matter Scope (iii)* allows for prediction, explanation and treatment as FOPs certain projects that neither of the current FOP objects proposals can accommodate. In fact, by taking *generics* as a case study, Part II of this dissertation exemplifies the *Widest Subject Matter Scope (iii)*, by introducing *speech acts* as FOP objects.

2.2.3 The Unifying Subject Matter Behind the Unity Assumption: Being Philosophical

Given the proposed most plausible unifying subject matter scope for FOPs, in this section, I suggest to resolve the *unifying subject matter substantive dispute* behind UA by settling on *being philosophical* as the unifying *subject matter category* behind UA. In more detail, my proposal states that being philosophical is the subject matter for FOPs whenever the suggested FOP object counts as *philosophical* (in addition to sharing the same unifying activity behind UA, see Section 2.3). In other words, lots of *objects* (e.g. concepts, terms, meanings, topics, reference, etc.) can count as appropriate *FOP objects*, but only if they can be classified as philosophical objects. I take philosophical objects to be the tools of philosophy, or tools used in philosophical engineering/analyzing. In a similar manner in which the method of chemical engineering/analysis is about engineering/analyzing *chemical objects* or objects that fall under the subject matter of chemistry, I believe that the method of philosophical engineering/analysis is engineering/analyzing philosophical objects or objects that fall under the subject matter of philosophy. This suggestion, however, opens an old metaphysical question about what counts as a philosophical object. In Chapter 4, Section 4.1, I offer my brief response.

2.3 Resolving the Unifying Activity Dispute

In this section, after considering some of the current proposals (Section 2.3.1) and suggesting the

most plausible unifying activity scope for FOPs (Section 2.3.2), I suggest how to resolve the unifying activity substantive dispute about UA (Section 2.3.3).

2.3.1 Current proposals

Despite acknowledging (see e.g. Burgess, Cappelen and Plunkett 2020; Scharp 2020; Thomasson 2020) the need to resolve the challenge about what constitutes the activity behind 'engineering,' discussion about the *unifying activity* behind UA has so far received less attention than discussion about the unifying subject matter behind UA. As Isaac (2020, 14) remarks: "Very few things have been said on the concept of 'engineering,' in the case of (conceptual) cognition, and for the purposes of conceptual engineering, except that it somehow involves an intentional and explicit design (e.g. Brun 2017: 2, 26 [see also Brun 2016: 1233 note 41])" (Isaac 2020, 14). More generally, independently of what one takes to be the subject matter behind UA, very few things have been said about the unifying activity behind UA.

Some of the terminology used for FOPs (meta-)theories got at least part of its labels depending on the aspects of unifying activity that it puts in the focus. For example, *conceptual ethics* stresses the underlying normative considerations of the engineering project; *ameliorative projects* stress the implementation stage, *replacement engineering* and *revisionary engineering* stresses certain preferable strategies as a result of the strategic planning stage, *explication projects* focus on explication of a certain class of representational devices, etc. For the purposes of this dissertation, I consider them as (at least partial) accounts of the unifying activity behind FOPs. Below is a brief recap of some of the (rather partial) proposals about the unifying activity behind FOPs.

Cappelen's (2018) Prudential Argument emphasizes *investigating defectiveness* and *amelioration* as two core unifying activities for those engaged in FOPs: "*The Prudential Argument*: If our representational devices can be defective in ways W1...Wn, then we should be engaged in two kinds of activities: *(i) investigating whether their concepts are defective and (ii) if defects are found, then ameliorating those concepts*" (Cappelen 2018, 40; italics mine).

Similarly, Scharp's (2020) radical therapeutic program stresses the following two activities: *identifying the problem* and *offering a solution*. In particular, he claims that "*it identifies our problem*, which is that most or all of our core concepts are inconsistent. And it *offers a solution*: use conceptual engineering to change our conceptual scheme so that we have concepts that work for us rather than concepts that tangle our thinking, confuse our beliefs, and interfere with our plans" (Scharp 2020, 441).

On the other hand, Isaac (manuscript), in what he calls "'Carnapian explication rebooted' for the purposes of conceptual engineering," sees the unifying activity as a three-step process including *conceptual regimentation, assessment phase* and *improvement phase*. According to Isaac, the first step or *conceptual regimentation* is conceived as a descriptive mode. The second step or *assessment phase* consists

in assessing target concepts with respect to (Carnapian) *fruitfulness*, *exactness* and *simplicity* as the three most important parameters, whereas the third step he takes to be the *improvement phase* which relies on Carnapian explication (see Issac manuscript).

Chalmers (2020) also distinguishes between three broad stages: *design stage*, *implementation stage* and *evaluation stage*:

> There's the design stage, where we design concepts. There are various ways to do that. One classic way would be to give a definition, or maybe an inferential role, or some paradigm cases, or something like that. Next is the implementation stage, like the stage where you actually build the bridge. In the implementation stage you actually use a concept, and perhaps try to get others to use it too. This is what Herman Cappelen calls conceptual activism. And then there's the evaluation stage, which plays a central role in the conceptual ethics work by people like Alexis Burgess and David Plunkett. Here the key is the evaluation of how good these concepts are in themselves and for certain purposes, to see how well they play key roles. (Chalmers 2020, 2–3)

Greenough (manuscript) proposes a wider scope for the unifying activity he dubs '*The Five Phase Model of Engineering*.' His model includes the following phases:

> *Descriptive Phase:* This yields an account of how the structure behaves under a variety of conditions together with an account of why it does so.
> *Evaluation Phase I:* This provides an assessment as to whether the behaviour of structure under these conditions meets the required goals (or other relevant standards of evaluation). If it does, then no improvement is needed. If it does not (and improvement is deemed worthwhile) then proceed to the next phase.
> *Design Phase:* This phase produces a new, hopefully improved design. If relevant, the design is suitably tested and test-data produced.
> *Evaluation Phase II:* This not only establishes whether and to what extent the proposed design really does meet the required goals, but also seeks to establish whether the proposed design can and should be implemented (which involves considerations of cost, need, and feasibility).
> *Execution Phase:* This phase then involves taking the design from drawing-board and test-lab to manufacture, or build, or propagation, or implementation. (see Greenough manuscript)

To date, even though their emphasis is on conceptual ethics, Burgess and Plunkett (2020), in their recent paper "On the relation between conceptual engineering and conceptual ethics," offer the most exhaustive account that focuses on the unifying *activity* behind UA. They stress three activities in particular: conceptual ethics, conceptual innovation, and conceptual implementation.

> Conceptual engineering is indeed definable in terms of these three activities (conceptual ethics, conceptual innovation, and conceptual implementation). But, we do not think it is so definable in a simple Boolean way. Rather, we think the term 'conceptual engineering' is best used to pick out something that is made up of mereologically complex activities whose parts fall into the categories associated with each of these three different activities. (Burgess and Plunkett 2020, 6)

However, I believe that the model for the unifying activity behind UA should withstand the test of

further refining and widening the scope. Recall, in Section 2.2.2.2, I offered a brief rationale for assuming that *the widest scope,* both in the case of the unifying subject matter and the unifying activity behind UA, *is the most plausible*. In the next section, I argue for the most plausible unifying activity scope behind UA, which I dub the '*Five Stage Recursive Model.'*

2.3.2 The Most Plausible Unifying Activity Scope: The Five Stage Recursive Model

In this section, I introduce and argue for the *Five Stage Recursive Model* as the most plausible *unifying activity scope* behind UA. The proposed model comprises the following stages: (i) *Identification*; (ii) *Evaluation*; (iii) *Strategic Planning*; (iv) *Implementation*; (v) *Re-evaluation*. As a restriction on the activity scope, i.e. on what can fall under the *FOP recursive stage*, I utilize both the *bottom-up (descriptive) approach,* by taking into account the commonalities among the current theorizing about the stages of the unifying activity behind UA as data points (see Section 2.3.1). Furthermore, I also utilize the *top-down (normative) approach*, namely I spell out what the *stage* of the unifying activity behind UA *should be,* by introducing and supplementing certain *stages* that have not (yet explicitly) emerged from the bottom-up approach. In Section 2.3.3, I then argue for *engineering* as the unifying activity *category*.

2.3.2.1. Identification

Identification is a largely descriptive stage in which the FOP object is being identified. With some exceptions (e.g. Scharp 2020 or Isaac manuscript), identification of a philosophical object has so far been either largely overlooked or considered merely *implicitly.*[60]

Even though there are other (potentially promising, yet understudied in the context of FOPs) descriptive approaches to identifying FOP objects (such as *historical* (see Burgess and Plunkett 2013b; Dutilh Novaes 2016; Plunkett 2016), *systematic* (see Brun 2016; Burgess and Plunkett 2013b; Thomasson 2020), or *experimental* (see Andow 2020; Fisher 2015; Koch 2019; Horvath and Koch 2020; Machery 2017; Shepard and Justus 2015)), this stage has so far been mainly perceived and practiced (often implicitly) as a version of philosophical analysis.[61]

For example, Scharp sees conceptual engineering as supplementing conceptual analysis. He claims that philosophy is "for the most part the study of what have turned out to be inconsistent concepts"

---

[60] The identification stage, as described here, might be more plausible for *(re-)engineering* projects (see Brun 2017) than *de novo* (see Chalmers 2020) or *innovation* (see Simion and Kelp 2020) projects.

[61] Philosophical analysis has traditionally been contrasted with philosophical naturalism (see Tye 1992, Devitt 1996, Kornblith 2002, Papineau 2020). Philosophical naturalists aim to arrive at the truth via *a posteriori* methods, as opposed to *a priori* armchair methods of philosophical analysis (see Quine (1951) for criticism of analyticity, purely armchair methods and the continuity between science and philosophy). However, there are also those who consider philosophical naturalism as part of philosophical analysis. For example, Halsanger (2006) argues that *descriptive* projects can be described as *a posteriori* method that includes philosophical naturalism or analyzing natural kinds as well as the analysis of social kinds in the similarly scientific manner. Furthermore, in contrast to Horvath (2017), Haslanger takes semantic externalism to be closely related to descriptive projects which she takes to be one kind of philosophical analysis (see Haslanger 2006, 107).

concepts" (Scharp 2020, 397). In *Replacing Truth*, Scharp (2013) offers an analysis of TRUTH yet he arrives at a contradiction, i.e. he shows that TRUTH is inconsistent. Even though Scharp (2013) gives an analysis that leads to inconsistency one may still argue that he, nevertheless, engages in an analysis. It is worth noticing that an analysis that leads to inconsistency is a different outcome than an analysis which is not possible. As another example, consider Haslanger's (2006, 2012) analysis of RACE and GENDER that she provides as part of a descriptive stage of her ameliorative analysis. In contrast to Scharp, Haslanger believes that she offered a true analysis of the concepts in question. However, she believes that the results of the analysis suggest that our concepts are deficient in a way that is connected to their semantic value only indirectly, i.e. the fact that RACE and GENDER have semantic value they have has bad (e.g. social, moral, and political) consequences (see Cappelen 2018).

Descriptive analysis has long enjoyed its primacy as a philosophical method. For example, some of those who acknowledge the existence of FOPs, see FOPs as belonging to the *non-descriptive* stage of philosophical analysis. For example, Horvath (2017) assumes "that philosophical analysis[62] is primarily a descriptive enterprise, and that explication should only be considered as a last resort—or at least as something that enters at a fairly late stage in the analysis of a philosophical category." (Horvath 2017, 3) My model, however, operates on the assumption that descriptive analysis can be part of the unifying activity behind UA.[63]

An objection to the assumption that descriptive analysis can be part of the unifying activity behind UA comes in a form of what Cappelen (2018) calls the 'Anti-Descriptive Argument' that claims that methodology behind UA does not have descriptive elements. Cappelen's (2018) objection is that:

> If your aim is to think about and understand some important philosophical phenomenon—say knowledge, causation or freedom—you have to figure out how best to think about those phenomena. (…) This kind of inquiry is essentially a normative enterprise. It asks how best to represent those phenomena and what might be defective about current ways of representing them. The assessment and improvement of concepts is at the core of philosophical practice, no matter what the topic. Your goal *cannot* be purely descriptive if you accept the Prudential Argument[64]—at the core of all philosophical activity is the continuous assessment of representational devices. (Cappelen 2018, 47–48)

Cappelen then imagines his opponent offering the following response to the above mentioned

---

[62] Horvath (2017) praises descriptive analysis because it tends to accommodate all the relevant counterexamples and does not easily accept stipulations in cases of disagreement. In particular, he believes that "… if explicative analyses were generally accepted in epistemology, then the debate about the analysis of knowledge would probably have taken a very different turn. For if one aims at an explicative analysis, one can stop taking counterexamples seriously much earlier and complete one's analysis based on other criteria, such as simplicity or theoretical fruitfulness (see also Weatherson 2003)." (Horvath 2017, 3)

[63] This may be a sign of yet another merely verbal dispute between Horvath's view and mine, in which case one might consider to call the whole thing something else than 'analysis.' I do not pursue this line of thought here.

[64] "*The Prudential Argument*: If our representational devices can be defective in ways W1...Wn, then we should be engaged in two kinds of activities: (i) investigating whether their concepts are defective and (ii) if defects are found, then ameliorating those concepts." (Cappelen 2018, 40)

Anti-Descriptive Argument: "Doesn't the Anti-Descriptive Argument assume that we first engage in important *descriptive* work: figuring out whether our concepts are defective and how they can be improved? Surely that's a descriptive task, isn't it?" (Cappelen 2018, 48). Cappelen offers the following answer:

> This reply shows a lack of understanding of the scope of the revisionary attitude which is motivated by the Prudential Argument. That argument *has no limit: it applies equally to the terms that occur in the 'Reply to the Anti-Descriptive Argument' above.* In other words, all the concepts involved in describing the critical/constructive project of conceptual engineering should themselves be subject to constant critical assessment and skepticism. In particular, we will need to assess the following concepts: 'concept', 'conceptual defect', 'descriptive work', etc. The very terminology in which you engage in the critical project is itself suspect. Once you endorse the Prudential Argument, there are no safe spaces. (Of course, even this articulation of the self-reflective nature of the core revisionist argument should be subject to that very same kind of criticism.) (Cappelen 2018, 48)

However, the tension that Cappelen (2018) induces can be resolved. According to the Five Stage Recursive Model for the method behind UA I propose, adopting identification as a descriptive phase in a recursive model does not need to diminish the importance of other stages of the main activity behind UA. In fact, in his later paper, Cappelen (2020, 146) too admits that there is no inconsistency or incompatibility between descriptive and revisionary projects and that they can be seen as complementary. Even though he contends that there is at least a 'loose tension' between the descriptive and revisionary projects, he admits that: "none of this is to say that these differences in goals, priorities and direction imply that the ameliorator shouldn't in part be guided and restrained by descriptive insights" (Cappelen 2020, 147).

Finally, this also opens a question whether philosophical analysis can be considered as a self-standing method or only as one of the recursive stages of the philosophical method behind UA. My tentative answer is that philosophical analysis functions as one stage of the philosophical method behind UA. However, sometimes one may be either stuck in the first stage because of, for example, not being aware of the possibility to engage in the second stage, or because of being satisfied with the current state of the philosophical object, often after already finishing one of the recursive loops of the model. Addressing these issues here in any more depth goes beyond the scope of this chapter.

2.3.2.2 Evaluation

Evaluation is a normative stage in which the FOP object is being evaluated. Those engaged in the unifying activity behind UA usually have different intellectual attitudes towards the quality of philosophical objects. As a useful heuristic, Cappelen (2018, 5–6) divides them in two main groups: *the*

representationally complacent vs. *the representational skeptics*.[65] According to him, they "represent a continuum, and the same person might be complacent in some domains, but skeptical in others" (Cappelen 2018, 6).

Furthermore, the quality of FOP objects can have profound effects. This is arguably one of the main reasons why evaluation as a *normative stage* has been widely recognized as an important element of the unifying activity behind UA (e.g. Cappelen 2017, 2018; Eklund 2015; Prinzing 2018; Scharp 2020; Simion 2018). For instance, some have argued that, in the case of representational apparatus, its quality "partly, and yet crucially, determines the quality of one's correlated cognitive activity" (see Isaac manuscript). In addition to that, some have also stressed the so-called 'wordliness' of the unifying activity behind UA, namely that FOP can have different effects on world and, thus, change the world itself (see Cappelen 2018; Haslanger 2012). In this section, I endorse and illustrate four underdeveloped axes of evaluation: *pluralism about the object of evaluation*, i.e. my model allows for evaluation of different FOP objects; *pluralist standards of evaluation*, i.e. my model relies on a broad reading of a FOP object value that allows for both semantic and non-semantic evaluation; *pluralism about the valence of evaluation*, i.e. my model includes positive, neutral and negative valence as a target of evaluation; *local and global evaluation*, i.e. my model distinguishes between FOP being evaluated at the local level (e.g. when applied to the one speaker or a group of speakers) vs. global level (e.g. when applied to the whole linguistic community).

Pluralist Standards of Evaluation

UA operates under the assumption that FOP objects can be defective in various ways (e.g. see Burgess, Cappelen, Plunkett 2020; Cappelen 2018; Scharp 2013, 2020). In principle, one can endorse two broad standards of evaluation. On the one hand, one can commit to one unique standard of evaluation. Isaac (2020, 7) calls this *'monist'* or *'absolutist' valuation*. For example, *scientific rationality* in the case of cognitive efficacy (e.g. Scharp 2020) is one candidate of such evaluation. On the other hand, one can commit to many different standards of evaluation. Isaac (2020, 7) calls this *'pluralist'* or *'relativist' valuation*. For example, in the case of cognitive efficacy, some have argued for different 'measures of appropriateness' (Sundell 2012, 757) which are usually taken to be context-dependent (e.g. Burgess and Plunkett 2013b; Nado 2019; Thomasson 2020) (cf. Isaac 2020, manuscript).

---

[65] "*The representationally complacent:* The representationally complacent uncritically take over the representational devices that are handed to them. They do their thinking and talking with whatever concepts they have inherited from their peers, teachers, and community. Someone who is representationally complacent can be a skeptic in the traditional epistemic sense: when she hears an assertion, she can be disposed to provide counter evidence and opposition, but she does so using only inherited concepts. More generally, the representationally complacent think and speak without a great deal of meta-reflection on the tools of their thinking and speaking.

The representational skeptics: Representational skeptics do not uncritically take over the representational devices handed to them. A significant part of their intellectual efforts consists in questioning and trying to improve the concepts of their peers, teachers, and community. When a representational skeptic starts reflecting on an issue, the first question she asks herself is whether the language used to articulate the key questions is good enough. The representational skeptic does not throw herself headlong into efforts to answer questions; rather, she first questions the concepts used to articulate the questions." (Cappelen 2018, 5–6)

The model I develop adopts a pluralist standard of valuation. A precursor to the pluralist standard can be found already in Carnap (1950, 4),[66] and, moreover, in Quine (1960, 258–259)[67] who, in *Word and Object*, also stresses the presence of "our interests and purposes." Cappelen (2020) follows a similar 'pluralist' or 'relativist' approach to valuation. He claims that "there is no unique correct explication of any term, the improvement is relative to contextually specific purposes. With that in mind, there is no reason why there should be a fixed set of theoretical virtues that are used to measure improvement. In certain contexts, non-theoretical virtues/advantages could make a big difference" (Cappelen 2020, 137–138).

Furthermore, Isaac (2020, 7–8) offers additional point of distinction when it comes to the pluralist standard of evaluation. His distinction takes into consideration whether FOP objects: *are philosophical or philosophically interesting* (Eklund 2015; Scharp 2020), *are deficient at large*, whether philosophically or not (Brun 2017; Cappelen 2018; Dutilh Novaes 2018; Scharp and Shapiro 2017), belong to *'open-texture' concepts* (Tanswell 2017), *help solving open questions* (Floridi 2011), *can be any concept whatsoever*, independently of its deficiencies (Simion 2018) and without any limitations such as 'bedrock concepts' or 'conceptual fixed points' (see Chalmers 2011; Eklund 2015).

I want to focus on yet another point of distinction which complies with the pluralist standard of evaluation. In addition to the standard *semantic valuation* which amounts to assessment of those FOP objects that are either directly or indirectly connected to the semantic aspect of FOP object, I also endorse a *non-semantic valuation* which stresses the assessment of FOP objects that are not (or cannot be) connected to any semantic aspect. Next, I illustrate some possible dimensions of semantic and non-semantic evaluation by utilizing some of the current rudimentary taxonomies of philosophical evaluation such as Scharp's (2013) cases of inconsistency; Cappelen's (2018, 2020) varieties of conceptual deficiencies; Greenough's (2017) conceptual malfunctions; Machery's (2017) conceptual invalidity; Isaac's (2020) theoretical and practical deficiencies.

(i) Semantic Evaluation

Direct Semantic Evaluation. Until recently philosophers have mainly been concerned with what can be classified as evaluation of FOP object deficiencies such as contradiction, inconsistency, conceptual

[66] "…if a solution for a problem of explication is proposed, we cannot decide in an exact way whether it is right or wrong. Strictly speaking, the question whether the solution is right or wrong makes no good sense because there is no clear-cut answer. The question should rather be whether the proposed solution is satisfactory, whether it is more satisfactory than another one, and the like." (Carnap 1950, 4)

[67] "We do not claim synonymy. We do not claim to make clear and explicit what the users of the unclear expression had unconsciously in mind all along. We do not expose hidden meanings, as the words 'analysis' and 'explication' would suggest; we supply lacks. We fix on the particular functions of the unclear expression that make it worth troubling about, and then devise a substitute, clear and couched in terms to our liking, that fills those functions. *Beyond those conditions of partial agreement, dictated by our interests and purposes, any traits of the explicans come under the head of "don't-cares"* (§38). Under this head we are free to allow the explicans." (Quine 1960, 258–9)

confusion, conceptual confission[68], incoherence, or vagueness[69] that are directly connected to semantic value (e.g. truth, reference, meaning, compositionally) (see Cappelen 2018, 33–34; Eklund 2002; Scharp 2013).

Indirect Semantic Evaluation. Evaluation can also amount to evaluation of FOP object deficiencies connected to its semantic value indirectly. As part of his taxonomy of conceptual deficiencies, Cappelen (2018) considers a deficiency according to which *semantic value* itself is not defective, i.e. a particular expression "has a [non-deficient] semantic value, but for that particular expression to have that semantic value has bad effects" (Cappelen 2018, 34).[70] He dubs this deficiency 'objectionable effects of the semantic value' (Cappelen 2018, 34). Cappelen (2018, 33–34) further distinguishes between different objectionable effects of the semantic value: *morally, politically, or socially objectionable effects*, *cognitive effects*, and *effects on theorizing*. For example, consider non-deficient semantic value of e.g. MARRIAGE, RAPE which can nevertheless bring about moral, social and political injustice when e.g. excluding same sex couples to be married, or not recognizing unwanted sex within marriage as rape (see Cappelen 2018, 28). Furthermore, according to Cappelen's interpretation of Leslie (2017), the use of certain kinds of expressions (e.g. generics) has negative cognitive effects on those using those expressions (cf. Chapter 6). Finally, according to Cappelen's (2018) interpretation of Clark and Chalmers (1998), our current concept BELIEF has bad effects on theorizing since even though concept BELIEF is not semantically deficient, its semantic value is constitutive of some other (theoretical) facts and that can bring about unwanted effects, such as preventing us from giving a unified theory of mind.

(ii) Non-semantic Evaluation

One can also evaluate FOP objects that are neither directly nor indirectly about any semantic aspect. For example, Greenough (2017) criticizes Scharp (2013) for taking philosophy to be mainly concerned with conceptual inconsistency which is, according to Greenough, only "one source of

---

[68] "However, there are many other kinds of conceptual engineering besides explication. One can identify a conceptual confusion, where someone assumes that one coherent concept can do a certain job, but it actually requires two or more (e.g., the concept of mass as it occurs in Newtonian mechanics, which I discuss in Chapter 2, is confused). Conversely, there can be cases where we assume that two distinct concepts can do two distinct jobs, but it turns out that these jobs are interrelated in an unforeseen way, which requires a single coherent concept; there is no name for this phenomenon, but we might call it conceptual confission (e.g., the concepts of space and time as they occur in Newtonian mechanics get replaced by a single concept, spacetime). And there are others as well." (Scharp 2013, 4)

[69] For those who do not advocate epistemicism about vagueness (e.g. Frege and Carnap) vagueness is a semantic deficiency. However, for an epistemicist about vagueness (e.g. Williamson) who does not take vagueness as a semantic phenomenon but rather as an epistemic phenomenon, vagueness falls under the non-semantic evaluation.

[70] Cappelen (2018, 34) offers two kinds of justifications for why some linguistic expressions (e.g. 'marriage' if it has an extension that excludes same-sex couples) can have bad effects:

> *Metaphysical justification:* the extensions of terms are in some cases constitutive of reality and so insofar as it matters what our society is like it will matter what extensions our terms have (for more on this idea, see Chapter 3, Sections 3.3.1 and 3.3.2).
>
> *Non-metaphysical justification*: Even if you don't think meanings of words are constitutive of social reality, you can think that as a matter of empirical fact the divisions and classifications we make will have very significant social effects. (Cappelen 2018, 34)

conceptual malfunction" and "perhaps not the most prevalent or interesting source" (Greenough 2017, 6). Even though Greenough (2017) focuses primarily on concepts as FOP objects and does not offer an elaborated taxonomy of conceptual deficiencies, he considers a great variety of deficiencies that may apply to different FOP objects. In particular, he (2017, 6) mentions the following conceptual malfunctions: *too parochial*, *too elitist*, *redundant* or *not fit to feature in any useful explanation*, *superseded, hackneyed*, or *systematically misapplied*, *loaded with ideological baggag*e or *serve as ongoing devices for deceit, discrimination, or oppression*.[71] Notice, however, how, since being under-characterized, some of these malfunctions can be interpreted ambiguously as belonging to semantic and/or non-semantic type of evaluation.

The above taxonomy is, however, not exhaustive. It could be expanded further depending on different factors, such as different values and aims of philosophical inquiry. I offer three quick examples of extensions of these general categories. First, in the case of *direct semantic evaluation*, instead of concept, one could take as an object of evaluation a linguistic expression such as e.g. the proposition "The present King of France is bald" and evaluate it with respect to its semantic properties, e.g. truth, reference, meaning and compositionally. Second, one could expand *indirect semantic evaluation* to include other effects than those suggested by Cappelen (2018). For instance, one could consider *psychological* and *emotional effects of the semantic value* on those using the expressions (see *Alice and the Mouse Case* below), or *effects of surprise (*e.g. consider a surprise meal called "Lucky 13 Mozza Dare" with the accompanying description of pizza content: "Trust in chef's creativity and try this delicious mystery surprise! Don't ask your server as even they don't know."). Third, within the *non-semantic evaluation,*[72] one could consider evaluating non-semantic dimensions of certain FOP objects such as their: *pragmatic dimension* (e.g. implicatures, assertability), *lexical dimension* (e.g. lexical effects), *epistemic dimension* (e.g. effects of knowledge or ignorance about the semantic value or content), *propositional attitude dimension* (e.g. belief, disbelief, hope, desire, etc.), *ontological* (e.g. empirical invalidity), etc.

Furthermore, it is important to notice that, depending on one's goals, one might want to evaluate multiple dimensions of the same FOP object. For instance, one could take a generic statement such as "Muslims are terrorists" and evaluate it with respect to its *semantic* properties, e.g. whether it is true or false. However, one could also evaluate it with respect to its *pragmatic* properties, e.g. what it is used to implicate when uttered in certain context, or with respect to *social, moral or political effects* that it might

---

[71] "Language might also do more; it might reinforce, cause, or constitute oppression. For instance, some theorists hold that slurs are words that dehumanize and demean (e.g., Jeshion 2013). On such views, uses of slurs do not merely reflect that some groups are oppressed, they also partially cause or constitute harm." (Ritchie 2019, 33).

[72] For example, consider different interpretations of the sentence "The present King of France is bald" discussed in the literature. According to Russell, the proposition expressed by this sentence is false, because there is no present king of France. According to Strawson, the speaker's utterance is defective because the speaker has simply failed to refer to anything and so has failed to make a complete statement (see Lycan 1999, 20).

contribute to (see Chapter 6).

Finally, it is also worth keeping in mind possible interconnections between: (i) various dimensions of different FOP objects, and (ii) various dimensions of the same FOP object. This should be taken into consideration during evaluation as well as in other stages since changes in value of one dimension could invoke changes in value of other dimensions. Consider the following two examples. First, one may argue that there is an interconnection between different FOP objects such as terms, concepts, arguments, theses and theories. For example, in some cases coming up with better concepts can improve our arguments, theses and/or theories. Or we could have a good argument but we could use deficient concepts in formulating our arguments, theses and/or theories. By improving concepts, we could, consequently, improve our argument, thesis or theory, and *vice versa*. In other words, we might sometimes be interested in evaluation of concepts because of their role in formulating useful theses, arguments or theories,[73] and *vice versa*. Second, consider an interconnection between theories, concepts and linguistic expressions tied to Sawyer's (2018) view about concepts and linguistic expressions. In a nutshell, according to Sawyer, concepts are stable and already in perfect order, i.e. there is no semantic deficiency. Instead, amelioration concerns level of theorizing, i.e. it can be seen as taking *theory* as a FOP object. According to her view, then, we could say that as part of an epistemic evaluation, we are evaluating how much epistemic progress we have made in developing better theories about our concepts.

Pluralism about the Valence of Evaluation

So far, among those who submit to UA, most of the theorizing about the evaluation stage has gone into identifying *negative valence*, the so-called *deficiency, harmness,* or *malfunction,* after which the aim has been to reach *positive valence* through certain action, the so-called *amelioration, beneficence,* or *improvement*. However, I believe that it is important to allow for the FOP object evaluation to include not only *negative*, but also *positive* and *neutral valence*.[74] Here are two reasons. First, since evaluation is a valuable guide to strategic planning in determining whether and what kind of strategy is required, it is good to be able to have a taxonomy that discriminates between different valencies. For instance, before one can consider if and what kind of action should take place it is useful to establish the current value and, thus, know what it requires to have positive, negative or neutral valence. Second, and equally important, because of a possibility of an interaction between different FOP objects, there might be need to deliberately make one FOP object worse in order to make another one better (e.g. the one more important at the time). Compare this with building a bridge which is very sturdy and safe but not very

[73] Cf. Chalmers 2020.

[74] Instead of being a sharp tripartite distinction, this distinction can be seen as coming in degrees or on a scale.

elegant for the purposes of getting as many people across as safely as possible in the case of an earthquake, as opposed to a bridge which is very elegant but not very sturdy and safe for the purposes of an art competition.

Furthermore, much of what I claimed about both semantic and non-semantic deficiency translates to neutralness and advantage. Here are some illustrations. In the case of *direct semantic evaluation*, one could look for FOP objects that are *coherent* instead of *incoherent*. Furthermore, when it comes to *indirect semantic evaluation*, instead of only looking for objectionable effects, one could also look for *morally, politically,* or *socially beneficial and neutral effects, beneficial and neutral cognitive effects, beneficial and neutral effects on theorizing.* For example, consider evaluation of beneficial effects of certain generics such as 'Girls are tough' or neutralizing effects in the case of appropriation of slurs (see Chapter 6). Within the *non-semantic evaluation*, one could, for example, evaluate a FOP object as being *too simple*, *too complex* or *too unspecific* leading to either negative, neutral or positive value depending on different considerations.

Local and Global Evaluation

Discriminating between local and global cases of evaluation can further bear a particular significance for local and global cases of strategic planning, implementation, and re-evaluation. Consider the following two cases. First, consider *Alice and the Mouse Case*. Terms such as 'cat' and 'dog' are usually considered as having neutral effects. However, during a conversation between Alice and the Mouse (from Carroll's *Alice in Wonderland*) the sole mentioning of words 'cat' and 'dog' caused a grave emotional distress to the Mouse. Alice, after realizing the emotional harm she has inflicted on the Mouse, decides to employ a strategy of referring to the animals in question by using their initials in the presence of the Mouse, hoping that these new labels will have diminishing or neutralizing effect, or that they would at least not cause any more harm to Mouse: "'You promised to tell me your history, you know,' said Alice, 'and why it is you hate—C and D,' she added in a whisper, half afraid that it would be offended again."[75] Let us now compare Alice's *local* strategy of abbreviation with a similar, yet *global* strategy that has been proposed in the case of certain pejoratives. Consider *The 'N' Word Case*. In this case, some have argued that, not only the *use* but also *mention* of a pejorative can have negative effects (see Williamson 2009 for use/mention distinction in relation to pejoratives; see Cappelen 2018 for lexical effects). In order to avoid or, at least, diminish the negative effects of pejoratives, some have proposed a neutralizing strategy which consists in introduction of 'N' word to be used both in the case of *use* and *mention* of the corresponding pejorative for African Americans. To summarize, in the first case, where terms 'cat' and 'dog' were not globally perceived as having negative effects so the neutralization was *local*. On the other hand, in the second case, given that perceived negative effects of 'N' word are *global*,

[75] It is interesting to observe that Alice is afraid that "the bad" material might "leak out" even if the abbreviation "C" and "D" is used.

so was the above mentioned proposed action for neutralizing those effects.

In the coming years, I expect from those theorizing and engaging in FOPs to start paying more attention to: (i) evaluating different FOP objects; (ii) systematizing and further expanding both semantic as well as non-semantic horns of evaluation, (iii) apart from negative, also paying attention to neutralizing and positive outcomes of evaluation.

2.3.2.3 Strategic Planning

Strategic planing is an argumentative stage during which, after engaging in deliberation, an optimal strategy for changing the valence of a FOP object is being designed and prescribed. Strategic planning aims at averting an issue prior to occurrence or to remedy the issue after it has already occurred. Which strategy is going to be deployed will be constrained by various factors. In this section, I briefly touch on some of these factors such *available strategies*; *normative and practical considerations*; *target valency for philosophical objects;* and *interactions between philosophical objects*.

Available Strategies

The four key axes I considered for the evaluation stage transfer to the strategic planning stage, i.e. (i) strategic planning takes different philosophical objects (as opposed to e.g. concepts only) as FOP objects, (ii) philosophical objects can have a broad variety of valuation, (iii) the target valence of a philosophical object can be positive, negative or neutral, (iv) strategic planning can apply at local and global level. With this in mind, I briefly outline five general strategies that may be applied to a particular philosophical object: *inaction, revision, introduction, abolition, replacement*.

Inaction is a strategy that may be considered not only when nothing else can be done but also when doing anything else will make things worse (*do no harm* principle).[76] For example, consider keeping TRUTH unchanged for everyday purposes (see Scharp 2013). Or, consider Chihara's (1979, 618) claim that "in the end, it may be wiser to live with the illness than to undergo the kind of surgery needed to remove all paradox-producing elements."

Revision, such as e.g. Carnapian explication, is a partial change of a certain philosophical object, typically constrained by the *limits of preservation*. Limits of preservation can concern e.g. the issues around the *topic continuity* (see Cappelen 2018) or conditions under which *changing the subject* is permissible (see Nado 2019 for radical functionalism).

Introduction may come in two modes: epistemic and metaphysical. The *epistemic mode* concerns introduction of an existing philosophical object, e.g. in order to reveal a new purpose for an existing

[76] See Scharp and Shapiro's (2017, 270) scenario of Pat's application for membership of the Secretary Liberation Club "where the solution is worse than the problem."

philosophical object or apply it to a new domain. The *metaphysical mode* concerns introduction of a new philosophical object into existence, e.g. consider inventing a new lexical item.

Abolition may also come in two modes: epistemic and metaphysical. The *epistemic mode* concerns *abandonment* which boils down to not using or applying an existing philosophical object, e.g. stop using it for a certain purpose or to stop applying it to a certain domain. The *metaphysical mode* concerns *eradication* which goes one step further by trying to get rid of the existence of a philosophical object in question. For example, consider a proposal to stop using certain generics, or an abolition of certain slurs and pejoratives. Or, consider the following case. In 2019, the duet 'Baby, It's Cold Outside,' written back in 1944, got banned from certain radio stations in the wake of the #MeToo movement. That came as the duet faced renewed scrutiny over what some say are inappropriate lyrics. One may argue that, in the later case, the object of engineering is at the level of discourse given that the whole text is being banned.[77]

Replacement is a strategy that may be employed where change is needed but revision is equally good or not a (better) solution, all things considered. Replacement may be seen as composed of an *abolition* of a certain philosophical object (either as abandonment or eradication) on the one hand, and as an *introduction* of another philosophical object (either an *existing one* or "*from scratch*") on the other hand. For example, consider local replacing TRUTH with ASCENDING TRUTH and DESCENDING TRUTH for the purposes of logic and mathematics (see Scharp 2013).

It is also worth noting that, when applied to a particular philosophical object, these strategies have been called different names such as 'concept broadening', 'concept narrowing', 'category extension', 'homonymous engineering' and 'heteronymous engineering,'[78] 'logocide' and 'semanticide,'[79] 'conceptual confusion' and 'conceptual confission,'[80] 'stipulative additions' and 'stipulative introductions,'[81] etc.

Here are some immediate questions for general strategies that those theorizing about this stage should eventually address.[82] *Inaction*: If there is a lack of change, can inaction count as an activity

---

[77] Furthermore, on the one hand, one can imagine that the meaning of the discourse is context sensitive, and that the song means different things in different contexts and times, e.g. as the social reality changes, the meaning changes from neutral to negative, too. On the other hand, one may argue that this song always had inappropriate meaning yet we revealed it only later, when certain conditions changed.

[78] See Chalmers 2020.

[79] For 'logocide' and 'semanticide' see Young (1991, Chapter. 4).

[80] See Scharp 2013, 4.

[81] See Deutsch 2020a, 14.

[82] The level of concern, among other things, may depend on the nature of a philosophical object in question.

behind UA? *Revision*: What are the limits[83] of revision?[84] *Introduction*: How is metaphysical introduction possible? Moreover, the epistemic mode of the introduction of philosophical objects raises the same worry: How did the already existing object come into existence in the first place? *Abolition*: How is eradication is possible? In what sense is abandonment possible? *Replacement*: What are the limits of replacement?

Normative and Practical Considerations

What determines whether a certain philosophical object will score negatively, neutrally or positively as well as what determines the course of strategic planning will also be largely governed by different *normative and practical considerations* (cf. Plunkett 2015). Here are some issues that those theorizing as well as those engaging in strategic planning may need to engage with.

Normative considerations can sometimes be conflicting. For example, McPherson and Plunkett (2020) describe the following normative considerations present in metalinguistic negotiation:

> Speakers in a metalinguistic negotiation might well express conflicting normative views about how a word should be used—views that will standardly be based on normative considerations about things *other* than words and concepts (e.g., how we should live, how we should organize our social/political institutions, or what objective joints there are in reality)—even if those views are expressed through pragmatic mechanisms (rather than in terms of literal semantic content). (McPherson and Plunkett 2020, 283)

There is also a deeper issue about potential existence of different normative considerations to pick from.[85] For instance, as Eklund (2017) notices, there might be alternative normative frameworks of normative concepts and if "there are alternative frameworks of normative concepts, there seems to be a question of, well, which one to employ. Which framework to employ has consequences for action" (Eklund 2017, viii). Furthermore, normative considerations can, at least in part, depend on other considerations, such as the facts about the aims an agent has. These can include different theoretical and practical considerations such as (e.g. scientific) aims, goals, interests, values, questions under discussion, knowledge, time limitation, available resources, pursuits over time, the debate in which FOP appears, current and target valency, nature of the philosophical object in question, context (see Machery 2017), different trade-offs (see Dutilh Novaes and Reck 2017; Nado 2019), etc. For instance, how one

---

[83] Cf. Nado's considerations: "Now, given this picture of the boundaries of revision, topic changes certainly seem to be on the menu—procedure x and procedure x′ might well delineate categories that are not plausibly the 'same thing'. Is this problematic? I'd argue that it is not. 'Changing the subject' isn't somehow inherently bad; if it is problematic, this must be because of its consequences. Strawson was concerned that Carnapian explication resulted in abandonment of the puzzles philosophers were concerned with; Cappelen's helpful generalization of this worry is that topic changes threaten to disrupt inquiry. A further worry Cappelen suggests for topic-changing revisions is that they produce verbal disputes and miscommunication." (Nado 2020, 17)

[84] They may partly depend on the type of the philosophical object as well on which similarity of the philosophical object we may want to preserve (e.g. topic (see Cappelen 2018), function (see Nado 2019)).

[85] See Plunkett and Sundell (2013) for further discussion. See also Thomasson (2016) and Ludlow (2014) for connected discussion.

should live or organize her social/political institutions can largely depended on change in different non-normative considerations. As an example, consider the change from positive/neutral to negative of the effects of the semantic value of the concept MARRIAGE prior to and after the 60's (i.e. before and after it included heterosexual couples).

Target Valency for Philosophical Objects

One of the main objectives of the strategic planning stage is to come up with a strategy for changing the valency of philosophical objects. The general assumption, among the theorists, has so far been that the outcome of the methodology of FOPs is the *improvement* or *amelioration* of a *deficient* philosophical object. In the previous section (see Section 2.3.2.2), I have argued that evaluation needs to be sensitive not only to negative valency of a philosophical object but also to the positive and neutral valency. Similarly, when it comes to proposing strategies, I endorse the view that the targeted outcome of the strategy can, apart from positive (*improvement*), also be negative (*harm*) or neutral (*neutralization*). In other words, one could utilize the methodology behind UA to make certain philosophical objects not only improved but also neutralized or deficient.[86] Below I consider three potential reasons for doing so, and endorse the first two. *Reason One*. Philosophical objects can interact between each other (see *Case Two* below). Thus, making a certain philosophical object (even more) deficient could be justified in certain cases when it would make a deficiency of some other philosophical object improved or neutralized. *Reason Two.* Philosophical objects can interact and affect non-philosophical objects. Thus, making a certain philosophical object (even more) deficient could be justified in certain cases when it would make a deficiency of some other non-philosophical object improved or neutralized (see Section 2.3.3 for a case of an endorsement of technically incorrect use of a term 'coronavirus' in order to achieve public health communication and social benefits). *Reason Three*. The methodology behind UA has the means to do it and, thus, one can do it, no justification required.

Finally, I want to tackle another under-investigated phenomena related to the strategic planning and valency of philosophical objects: *supererogation* and *subrogation of philosophical objects*,[87] i.e. the phenomena of improving beyond duty and harming without prohibition.

Supererogatory Action: Non-required Improving. Can there be an improvement of a certain philosophical object that is not required, and even if there were such improving, how come it is optional or beyond the call of duty? For instance, some philosophical objects might turn out to be good yet improvable. In some cases, we might want to consider to improve a philosophical object that is neutral.

---

[86] Notice that this proposal, potentially, goes against Simion's (2017, 10) *epistemic limiting procedure*: "A representational device should be ameliorated iff (1) There is all-things-considered reason to do so and (2) The amelioration does not translate into epistemic loss."

[87] For more on supererogation and suberogation see Heyd (2019).

Here are some potential cases that might be seen as going beyond the call of duty. For example, consider a case of *reappropriation*, *reclamation* or *resignification* of pejorative words or artifacts that went beyond its neutralization and has achieved the appraisal or even the hortatory level. The examples can be found in the areas of human sexuality, gender roles or sexual orientation (e.g. 'dyke,' 'faggot,' 'queer,' 'tongzhi'); politics (e.g. 'Cavalier,' 'Roundhead,' 'Yankee'); religion (e.g. 'Jesuit'); race, ethnicity, and nationality (e.g. Black, nigga, Curry, Jew, where trash, Dog), etc.[88]

Suberogatory Action: Non-prohibited Harming. Can there be harming coming from a certain philosophical object that is not prohibited, and even if there were such harming, how come it is not prohibited or how can such harming be permissive? In some cases, we might want to consider making a certain philosophical object deficient in order to e.g. raise or neutralize the value of some other philosophical object. For example, this could involve cases of deliberately not explicating certain vague concepts or statements because they might become cognitively too taxing. More generally, Scharp and Shapiro 2017 ask the following question: When is it reasonable to replace an inconsistent concept? They take the concept of truth as a case study and proceed under the assumption that: "It might turn out that the concept in question is still used in certain situations even by those who recognize that it is inconsistent" (Scharp and Shapiro 2017, 257).[89] For instance, consider Scharp's (2013) suggestion of a local engineering of a concept TRUTH, allowing for the less optimal use of a deficient concept TRUTH in everyday discourse.

Interactions between Philosophical Objects

A possible interconnection between philosophical objects also needs to be taken into consideration during strategic planning. Here are some possible cases.

Case One. Sometimes a certain philosophical object can be deficient at its *semantic level* which may in result produce certain bad effects (e.g. political, social, moral). However, in order to get rid of these bad effects, the proposed strategy may have to take place at e.g. the *pragmatic level* (because, for example, semantic change would be too hard to obtain or it may take too long and one needs the results as soon as possible).

Case Two. Sometimes, in order to ameliorate a particular philosophical object, we might want to suggest a certain strategy that would make another philosophical object deficient. For example, we

[88] As another class of examples, one can consider the tricky question of how meaning-assignments to languages relate to engineering in the case of Davidson and Lewis. For example, consider Davison's idea about the theory of meaning being governed by the principle of truth maximization (see Davidson 1973, 1974) which goes beyond the call of duty. Or, consider Lewis' (1984) idea that the referents of our terms should, as much as possible, be "the ones that respect the objective joints in nature" (Lewis 1984, 227).

[89] Scharp and Shapiro (2017, 264–272) spell out four conditions for replacement, i.e. Does C Figure in a Valuable Project?, Does the Inconsistency of C Inhibit the Pursuit of the Project?, Can the Trouble Brought by C Be Avoided?, Is a Replacement, or Batch of Replacements, for C Available?, which answer the question: "When Is the Cure to Be Preferred over the Disease?"

could sacrifice the exactness, scientific, or epistemological standards to improve the social, moral, or political standards. In other words, making meaning and reference deficient could sometimes improve the social, moral, or political effect of the sentence in which they occur.

Case Three. Sometimes changes in one philosophical object will go hand in hand with another philosophical object. For instance, one might come up with a very good argument but use deficient concepts in formulating her thesis. By improving concepts she could be in a better position to formulate useful thesis, or offer powerful arguments.

2.3.2.4 Implementation

Implementation is an actionable stage in which the strategy or the strategies for changing the valence of a FOP object are being tested and implement. It is often assumed that those defending the philosophical method behind the UA submit to a version of a *Naïve View of Implementation.* As a first pass, the view says that the method behind FOPs assumes a considerable degree of *feasibility*. Despite the optimism shared among those who endorse the Naïve View, some of the most pressing issues in debates around the implementation of the method behind UA concern its *feasibility*, *value* and *methodological newness or negligence*, i.e. the question whether engineering in philosophy is an old practice that has not yet been subject to systematic theorizing (see Deutsch 2020a). In particular, Deutsch (2020a) offers a rather pessimistic dilemma for the method behind UA. He argues that those advocating for the method behind UA are either "ignorant of how conceptual engineering can be implemented, or else it is trivial to implement but of very little value, representing no new or especially fruitful method of philosophizing" (Deutsch 2020a, 1). In Chapter 4 (Section 4.1.3), after discussing the main challenges stemming from Deutsch's pessimistic dilemma for the method behind UA I present the *hard problem of implementation* and show how my model squares with respect to this problem.

It is also worth mentioning that on some accounts, one might be able (to some extent) eschew the talk about implementation challenge. For example, the advocates of conceptual ethics (see Burgess and Plunkett 2013a,b); Plunkett 2015) seem to want to stay on the conservative side when it comes to implementation. As Plunkett (2015, 842) put it: "There is no reason, after all, that one can't advocate for a conservative view in conceptual ethics, rather than a reforming or revolutionary one" (Plunkett 2015, 842). Similarly, Koch stresses: "a difference between proposing theories whilst not having a recipe for how they can be implemented on the one hand, and proposing theories whilst knowing that they cannot be implemented on the other. The first is compatible with the idea that theories provide blueprints for action" (Koch 2018, 10). More explicitly, Greenough (manuscript) states: "The job of the philosopher is to isolate a philosophical problem, correctly identify the semantic defects which give rise to this problem, and to produce a kind of blueprint for semantic change (or replacement)—for how our

language should be (semantically) revised (or replaced) such that the problem does not arise in the revised language" (Greenough manuscript).

2.3.2.5 Re-evaluation

Re-evaluation is a hybrid descriptive-normative stage in which the strategy or the strategies for changing the valence of a FOP are being re-evaluated after being tested and implemented. It often comprises identification and evaluation of an implemented FOP strategy (see Section 2.3.2.1 and Section 2.3.2.2). It can be seen as the fifth stage of the first cycle or the first stage of any consecutive *recursive cycle* of a FOP.

I take the main roles of re-evaluation to be to:
(i) identify and describe the post-implemented philosophical object;
(ii) account for potential changes in evaluation approach;[90]
(iii) detect potential errors during the strategic planning stage;
(iv) detect potential errors during the implementation stage;
(v) test the feasibility of the proposed strategy and the feasibility of its implementation.[91]

Furthermore, re-evaluation can be understood as a corollary of Cappelen's (2020) Master Argument. According to The Master Argument, no matter what concept (or, more generally, what philosophical object) we end up with it is unlikely that it will be the most optimal concept (or, more generally, philosophical object). Thus, we need to critically assess our concepts (or, more generally, philosophical objects) (see Cappelen 2020, 134). Moreover, the re-evaluation stage may initiate recursion. This is compatible with the sentiment of the current literature about UA, especially with the strand that supports the idea of re-engineering, according to which even the concept CONCEPTUAL ENGINEERING can itself be subject to further re-engineering (see also Brun 2016, 2017; Cappelen 2018, Dutilh Novaes 2016, 2018; Dutilh Novaes and Reck 2017; Plunkett 2016).

With this I conclude the *bottom up* and *top-down* building of the model for the unifying activity behind UA. In particular, by taking into consideration the commonalities between FOPs as well as by taking into consideration (general and domain-specific) theories about FOPs as data points, I have sketched a blueprint for a *bottom up building* of the *five stages of the recursive model for the philosophical*

[90] Such changes may involve changes in purposes such as changes in certain moral or political goals, or changes that result from technological changes, or changes as a result of theoretical developments (see Cappelen 2020, 140).

[91] Testing can be seen as part of the re-evaluation given that it may be argued that testing already requires some kind of implementation.

method behind UA. Furthermore, by spelling out what the model of the unifying activity *should be* by introducing and supplementing certain stages and details that have not (explicitly) emerged from the bottom up building of the model, I have sketched a blueprint for a *top-down building* of the *five stages of the recursive model for the philosophical method behind UA.* Moreover, even though this model speaks to the nature of the unifying activity behind UA, it may, nevertheless, be seen as only one of the loops of the *recursive process* of building and improving the most plausible unifying activity scope behind UA. In the next section, given the proposed most plausible unifying activity scope for FOPs, I suggest to resolve the unifying activity substantive dispute behind UA by settling on *engineering* as the unifying activity behind UA.

2.3.3 The Unifying Activity Behind the Unity Assumption: Engineering

By taking into consideration currently proposed FOP objects as data points, as well as the significance of having the widest unifying activity scope for FOPs, next, I propose how to resolve the *substantive unifying activity dispute* behind UA. In particular, my thesis is that *engineering* should be considered as the unifying activity behind UA. My proposal relies on the premise that engineering is a unifying activity of the method behind UA whenever the suggested FOP complies with the *five stage recursive model* (in addition to operating on philosophical objects, see Section 2.2.3). In the rest of this section, by elucidating the two features of engineering continuity: *engineering resemblance*, *division of engineering labour,* I support the above thesis by pointing at two kinds of engineering continuity:[92] (i) continuity between engineering different philosophical objects; (ii) continuity between engineering philosophical objects and engineering other kinds of objects. Finally, I suggest that the synergy of different kinds of engineerings points to *engineering holism*.

Engineering Resemblance

Let me briefly motivate this view by considering a parallel view that philosophy is broadly continuous with the sciences. According to Williamson (2007):

> The unexceptional nature of philosophy is easier to discern if we avoid the philistine emphasis on a few natural sciences, often imagined in crudely stereotyped ways that marginalize the role of armchair methods in those sciences. Not all science is natural science. Whatever crude empiricists may say, mathematics is a science if anything is; it is done in an armchair if anything is. In no useful sense are mathematical questions conceptual questions. If mathematics is an armchair science, why not philosophy too? (Williamson 2007, 4)

[92] This section is about continuity of the unifying activity of FOPs. It is worth noting that there is a similar debate about the continuity of the unifying subject matter of FOPs, better known as Strawson's challenge (see Strawson 1963) to Carnapian explication (see Carnap 1963) which has gained a lot of traction in the current debate about engineering in philosophy. Different answers have been offered with respect to how and whether this continuity can be preserved (e.g. see Brigandt 2010; Brigandt and Rosario 2020; Cappelen 2018; Chalmers 2011; Jorem 2020; Koch 2020b; Nado 2019; Plunkett and Sundell 2013; Prinzing 2018; Simion and Kelp 2020; Thomasson 2020; Weinberg 2006).

Similarly, one could argue for an *anti-exceptionalism about engineering*, or the continuity between philosophical engineering and engineering in other sciences. As Pinder (manuscript) puts it:

> … there are borderline cases between philosophy and science within linguistics, cognitive science, evolutionary biology, foundational physics, pure mathematics, and so on. And there is significant methodological continuity between philosophy and the sciences: philosophers and scientists alike present arguments, build theories, provide explanations, make intuitive judgements, appeal to (and sometimes generate) empirical results, and so on. This is not to deny that there are significant differences between most of philosophy and, say, the experimental natural sciences; it is rather to locate those differences on a continuous spectrum. (Pinder, manuscript)

Furthermore, those who theorize about FOPs have already recognized in various ways a significant level of similarity between *engineering* philosophical objects and other kinds of engineerings. One of the first philosophers to draw this connection was Simon Blackburn (1999). In the introduction to *Think*, Blackburn (1999) introduces his work as doing conceptual engineering. He draws parallels between the activity of an engineer (in the 'real' sense of the word) and the activity of a philosopher. His focus of comparison is mainly on studying and understanding the structure of the object of engineering. As he puts it:

> I would prefer to introduce myself as doing conceptual engineering. For *just as* the engineer studies the structure of material things, so the philosopher studies the structure of thought. Understanding the structure involves seeing how parts function and how they interconnect. It means knowing what would happen for better or worse if changes were made. This is what we aim at when we investigate the structures that shape our view of the world. Our concepts or ideas form the mental housing in which we live. We may end up proud of the structures we have built. Or we may believe that they need dismantling and starting afresh. (Blackburn 1999, 1–2; italics mine)

More recently, others (see Isaac 2020; Greenough manuscript; Chalmers 2020) have also drawn attention to resemblance between the unifying activity behind UA and other kinds of engineerings. For example, Isaac (2020) assumes that there is an engineering process underlying his project: "At the background of this engineering project of the concept of '(cognitive) engineering' for the purposes of conceptual engineering is the assumption that cognition is an engineering process—namely, that of the informational relationship we enact, as cognitive agents, with our environment (Sect. 3.1.2)" (Isaac 2020, 16). Greenough (manuscript) also takes a non-metaphorical stance towards the unifying activity behind UA by claiming: "Question: Is Conceptual Engineering really a form of Engineering? Answer: Arguably yes! This is NOT a metaphor." Chalmers (2020, 2), draws a useful analogy between conceptual engineering and other kinds of engineerings. In particular, after offering parallels with mechanical and

software engineering,[93] he drives home the resemblance between engineering in philosophy with other kinds engineering by offering the following definition:

> On this definition, engineering is the process of utilizing knowledge and principles to design, build, and analyze objects. The key part is 'design, build, and analyze'. (…) That's not a bad definition, except that 'analyze concepts' already has a meaning which is not totally apropos in this context. Maybe 'evaluate concepts' is better. And maybe 'implementing' is better than 'building' where concepts are concerned. With these tweaks, we get the following definition: conceptual engineering is the process of designing, implementing, and evaluating concepts. (Chalmers 2020, 2)

I also take a strong non-metaphorical stance towards the unifying activity behind UA being *engineering of philosophical objects*. I believe that engaging in FOPs carries a *family resemblance* to other engineering projects outside philosophy, i.e. those operating on non-philosophical objects. Moreover, even if current engineering in philosophy does not resemble, I believe that it *should* resemble other kinds of engineering projects. The Five Stage Recursive Model I proposed as being the most plausible unifying activity scope for FOPs follows the general stages common to unifying activities of other kinds of engineering.

Here is an additional motivation for modeling my proposal on other engineerings. Historically, it has long been the case that philosophical methodology gave birth to other methodologies. However, nothing precludes the arrow of the methodology transfer to go the other direction, namely for philosophy to adopt the engineering methodology that has already been used in certain sciences. A similar transfer has already occurred among other sciences, and I believe that philosophy should not be an exception. For instance, mechanical, chemical, civil and electrical engineering were among the first to be recognized as forms of engineering. As engineering methodologies with different subject matters have been developed, the number of kinds and subkinds of engineering disciplines has significantly increased from the above mentioned ones to the new ones, including engineering different subject matters such as aerospace, aeronautics, environment, nuclear, petroleum, computer, software, biology, genetics, etc. Furthermore, it is also worth noticing that the (methodological) development may be contingent on the development of human and scientific thought and inquiry. Similarly, I believe, it is a contingent feature of the developments in philosophical methodology that only recently have we started engaging in a more systematic theorizing about engineering in philosophy.

Division of Engineering Labour

Many have already noted the importance of the *interaction between philosophical objects and reality*

[93] "You can see all three of these things playing a role in domains like bridge engineering. You design a bridge, you implement a bridge, you evaluate the bridge to see how well it's doing. If the evaluation isn't positive, you design some repairs and you implement the repairs. And so on. You also see something like this in software engineering. You design a program, implement the program, evaluate the program, and so on in a continuing circle." (Chalmers 2020, 3)

(e.g. Burgess and Plunkett 2013a; Cappelen 2018; Eklund 2015; Floridi 2011; Nado 2019; Prinzing 2018; Thomasson 2020; Haslanger 2000, 2020). In particular, as Haslanger (2000) stressed:

> …we begin by considering more fully the pragmatics of our talk employing the terms in question. What is the point of having these concepts? What *cognitive or practical task* do they (or should they) enable us to accomplish? Are they *effective tools to accomplish our (legitimate) purposes*; if not, what concepts would serve these purposes better? (Haslanger 2000, 33; italics mine)

Burgess and Plunkett (2013a, 1096–1097) draw attention to the interaction and enabling conditions of engineering different philosophical objects as well as its more indirect and non-philosophical consequences.

> [W]hat concepts we have fixes what thoughts we can think. [. . .] Arguably, our conceptual repertoire determines not only what beliefs we can have but also what hypotheses we can entertain, what desires we can form, what plans we can make on the basis of such mental states, and accordingly constrains what we can hope to accomplish in the world. Representation enables action, from the most sophisticated scientific research, to the most mundane household task. It influences our options within social/political institutions and even helps determine which institutions are so much as thinkable. Our social roles, in turn, help determine what kinds of people we can be, what sorts of lives we can lead. Conceptual choices and changes may be intrinsically interesting, but the clearest reason to care about them is just that their non-conceptual consequences are pervasive and profound. Burgess and Plunkett (2013a, 1096–1097)

Cappelen (2018) points to a connection between a continuity of revision in conceptual engineering and social reality:[94]

> If language is constitutive of social facts, then there is a corollary: the process of conceptual revision is also a process that involves a revision of the relevant social facts. There's a kind of dynamic evolution of the social facts that goes hand in hand with the dynamic evolution of the concepts. So conceptual engineering is engineering of the social world as much as it is of our concepts. (Cappelen 2018, 44)

Similarly, I believe that we *should* make room in philosophy for a highly integrative version of

---

[94] "According to a Searle-style view, our concept 'money' plays a constitutive role in both the creation and continued existence of money. Suppose then that these concepts evolve. Suppose we find that our current concept of money is defective; we revise and ameliorate the concept.
(…)
Many think that the institution of gender is a social fact. Suppose, to make things simple, that social fact is in part constituted by the terms 'man' and 'woman'. Suppose further that a Haslanger-style ameliorative project succeeds: it changes our gender concepts in the ways she advocates. We have then changed social reality: gender has been changed." Cappelen 2018, 44)

interdisciplinarity[95] between different kinds of engineerings,[96] both when it comes to engineering philosophical as well as non-philosophical objects. In particular, connections, both between engineering philosophical as well as non-philosophical objects, have also been recently made by some. For example, consider Thomasson's (2020, 454) claims that: "… it is an under-appreciated point that conceptual engineering, no less than civil, does not take place in a vacuum, and that it is extremely important to note and be responsive to the inferential connections between the term in question (whether one we are considering revising or eliminating) and our other terms and broader practices" (Thomasson 2020, 454). Even though Thomasson (2020) brings out an element of interconnection between different kinds of engineering with other practices, some of which could, arguably, be other kinds of engineerings.

On the other hand, Chalmers (2020) makes a more direct connection between interaction of different kinds of engineerings of philosophical objects:

> This immediately leads us to delicate questions about the connection between conceptual engineering and linguistic engineering. Sometimes these phrases are used almost interchangeably. Certainly it's the case that wherever you find conceptual engineering, you almost always find some linguistic engineering. Whenever there's a proposal about a new concept, there's also a proposal about a word for it to be attached to. (Chalmers 2020, 9)

I, thus, believe that this line of thought can be generalized beyond philosophical engineering. In particular, sometimes one kind of engineering can serve as an enabling condition or as means to an end for another kind of engineering project. Call this a *division of engineering labour.* On the one hand, when it comes to engineering of philosophical objects, one can argue that there is a possibility for a variety of interactions between different kinds of engineerings of philosophical objects, e.g. epistemic, semantic, pragmatic, or terminological engineering. For example, in some cases, engineering beliefs can come before engineering concepts (see Pollock 2020). On the other hand, one can argue that engineering of philosophical objects can also interact with and affect engineering of non-philosophical objects, as well as *vice versa*. For example, those who care about social engineering and social justice want to engineer better actions. One way to contribute to that is by *thinking* and *speaking* better with, for example, some help from e.g. linguistic, terminological and/or cognitive engineering. Or, in some cases, semantic and/

---

[95] Compare this with NASA's definition of 'systems engineering' that its defined as "a methodical, multi-disciplinary approach for the design, realization, technical management, operations, and retirement of a system. A "system" is the combination of elements that function together to produce the capability required to meet a need. The elements include all hardware, software, equipment, facilities, personnel, processes, and procedures needed for this purpose; that is, all things required to produce system-level results. The results include system-level qualities, properties, characteristics, functions, behavior, and performance. The value added by the system as a whole, beyond that contributed independently by the parts, is primarily created by the relationship among the parts; that is, how they are interconnected. It is a way of looking at the "big picture" when making technical decisions. It is a way of achieving stakeholder functional, physical, and operational performance requirements in the intended use environment over the planned life of the system within cost, schedule, and other constraints. It is a methodology that supports the containment of the life cycle cost of a system. In other words, systems engineering is a logical way of thinking."(National Aeronautics and Space Administration 2019)

[96] As a corollary of my view, philosophical engineering does not have a central role—on this view, there is no Philosopher King. Instead, philosophical engineering is only one kind of engineering alongside with other engineerings.

or conceptual engineering can affect software engineering, or conceptual engineering can affect genetic engineering. Furthermore, engineering of philosophical objects can happen or can be enabled by other kinds of engineerings. For instance, one could argue that mechanical or social engineering (by creating certain ontology), can enable conceptual and terminological engineering (for from world to content engineering cf. Greenough manuscript). For example, one could argue that the term 'robot' had function but it had an empty extension (outside of the world of fiction) until we created a certain ontology.

It is also worth noting that allowing different kinds of engineerings working together could contribute to resolving the so-called 'implementation challenge.' For example, sometimes the implementation stage of one kind of engineering (either in part or entirely) can be outsourced to another kind of engineering (see Chapter 6 for outsourcing of engineering generics). Given this potential to contribute to the implementation challenge, one could defend a stronger claim according to which, there *should* be a division of labour between different kinds of engineerings in order to help with the implementation challenge. For example, one could argue that in order to engineer a just society one should seek out contribution of engineering certain philosophical objects. Or, in order to have some progress in genetic engineering engineering of certain philosophical objects may be required. As a successful example from the past, consider a case of splitting the concepts MASS and WEIGHT. One could argue that it was necessary (or at least very helpful) to do some conceptual, semantic, and terminological engineering first, before we could have done some mechanical engineering of certain devices (that have proven to have great value for physics as well as for our everyday lives).

Finally, consider the following two *division of engineering labor* examples. Recent phenomena that go by the terms of 'coronavirus' and 'social distancing' can be seen as examples of *division of engineering labour* where behavioral engineering is aimed to be achieved *via* epistemic, terminological and conceptual engineering. Furthermore, these cases show how sometimes we might (have good reasons to) want one kind of engineering to *trump* another.

In the case of 'coronavirus,' we might allow for behavioral engineering to trump terminological, semantic and conceptual engineering. From e.g. public health communication perspective it might be more important to reach the desired behavior with technically incorrect or semantically 'deficient' term or concept which results in desired lexical and behavioral effects, because this can contribute towards engineering of behavior, as opposed to having a 'semantically correct term' with negative or no lexical or behavioral effects.

In the case of 'social distancing,' one could argue that, before the advances in our social and technological reality both the concept SOCIAL DISTANCING as well as the term 'social distancing' were apt for achieving the desired behavior. However, after the advances in our social and technological reality, one part of the phrase, namely 'social,' become less apt for our present purposes of social behavior

engineering. In more detail, in the past being physically present was a requirement for being social to a much larger degree than today. Thus, in the past, preventing people from the physical contact prevented them also from being social to a larger degree than today. Today, with the help of technology, it is possible to achieve the physical distance while still remaining social. Given that with the advances in our social and technological reality the overlap between the extension of PHYSICAL DISTANCING and SOCIAL DISTANCING became smaller, the concept SOCIAL DISTANCING as well as the term 'social distancing' became less apt for the role they play in the present day pandemic related behavior engineering. This may call for some conceptual and terminological engineering, and most probably, a shift in conceptual (and, possibly, even a semantic meaning) change, followed by a terminological change. For instance, a few months into the pandemic, the World Health Organization[97] in their official statements started using the term 'physical distancing' instead of 'social distancing.' It is also worth noting that, for the purposes of this particular behavioral change and psychological stability, both semantic as well as lexical effects of 'physical distancing' may be greater than those of 'social distancing' (cf. Chapter 3).

[97] See IFLScience 2020.

CHAPTER 3

TERMINOLOGICAL DISPUTE RESOLUTION OF THE UNITY ASSUMPTION

Abstract

In this chapter, I offer a novel resolution of the *terminological* dispute about the philosophical method behind the Unity Assumption (UA). First, I introduce a general definition of a *terminological dispute* and, as a case study, I suggest a terminological dispute about the philosophical method behind UA. Second, I evaluate two prominent terminological choices, i.e. 'conceptual engineering' (see Scharp 2013, 2020; Eklund 2015; Cappelen 2018; Burgess, Cappelen and Plunkett 2020; Chalmers 2020) and 'conceptual ethics' (see Burgess and Plunkett 2013a,b; Plunkett and Sundell 2013; Burgess 2014; Burgess and Plunkett 2020), against the two main conditions that I suggest one should take on board when making terminological choices for the philosophical method behind UA: (i) *the semantic correctness condition* drawing on Belleri (2018), and (ii) *the beneficial lexical effects condition* drawing on Cappelen (2018) and Chalmers (2011, 2020). Finally, I engage in a *terminological dispute resolution* by arguing that we should introduce a new term, 'philosophical engineering.' For the philosophical method in question this terminological choice, I argue, fares better against its competition with respect to both of the above mentioned conditions.

3.1 Terminological Disagreement

Our attitudes towards terminological choices fall along a spectrum from *intellectual complacency* to *intellectual skepticisim*.[98] On the one hand, the *terminologically complacent*, i.e. those who unquestioningly accept the terms that are handed to them, do not care about the words or terminology they use for a certain phenomenon. They believe that their terminological choice, when it comes to naming a particular e.g. concept, theory, argument, assumption, method, field, etc., is not all that relevant. Terminologically complacent people introduce new or utilize existing terms to which they attach (new) meanings/definitions without giving terminology any special attention. What is relevant, they believe, are the *meanings* of these terms and what the nature of the phenomenon represented *is*. As Chalmers (2011, 517) puts it: "Often, however, words do not matter. It often happens that we are concerned with a first-order domain, not with the usage of words, and in such a way that nothing crucial to the domain turns on the usage of words" (Chalmers 2011, 517).

On the other hand, *terminological skeptic*, i.e. those who do not unquestioningly accept the terms that are handed to them, believe that our terminological choices do matter. They pay attention to which terms serve our theoretical and practical purposes best and introduce new or utilize existing terms to which they attach (new) meanings/definitions. They choose their terms rather carefully and critically. For example, Cappelen (2018, 4) stresses the significance as well as difficulty of terminological choices: "This book is in part about the significance of terminological choices and so it's fitting that the terminology for talking about this field itself presents us with difficult choices" (Cappelen 2018, 4).

As opposed to the terminologically complacent, terminological skeptics are more susceptible to enter into a *terminological disagreement*. A terminological disagreement can be seen as closely related to the verbal dispute. The following passage from Chalmers (2011, 516–517) illustrates the importance, on the one hand, the interconnection of terms and meanings or senses connected to terms and, on the other hand, the interconnection of terms and things connected to terms that go beyond meanings or senses such as *connotations*, *associations*, and *attitudes* connected to terms.

> In cases where words have fixed *connotations and associations*, too, verbal issues often have serious practical import. This applies especially when those connotations are normative. What counts as 'torture' or as 'terrorism' might be, at one level, a verbal issue that a philosopher can resolve by distinguishing senses. But in a rhetorical or political context, words have power that transcends these distinctions. If the community counts an act as falling into the extension of 'torture' or 'terrorism', this may make a grave difference to our attitudes toward that act. As such, there may be a serious practical questions about what we ought to count as falling into the extension of these terms. (Chalmers 2011, 516–517; italics mine)

However, one should distinguish a terminological dispute from a verbal dispute given that a

[98] I model this on Cappelen's (2018, 5) useful heuristic about the "representationally complacent" and "representational skeptics."

terminological dispute does not need to boil down (only) to the meaning of the term. Drawing on the notions of *verbal dispute* and *metalinguistic negotiation* I further explicate a *descriptive* and *normative character* of a terminological disagreement. The *descriptive* character of a terminological disagreement I model on Chalmers' (2011) notion of a *verbal dispute*.[99] Call this a 'terminological dispute.'

> *Terminological Dispute*: A dispute over a term T is terminological when, for some T, the parties *do not disagree* about the meaning behind T, rather the dispute over T arises wholly in virtue of the disagreement about whether to use T rather than an alternative term, A.

The *normative* character of a terminological disagreement I model on Plunkett's (2015) and Plunkett and Sundell's (2013) notion of a *metalinguistic negotiation*.[100] Call this a 'terminological negotiation.'

> *Terminological Negotiation:* A terminological dispute that concerns a normative issue about what term we *should choose*, or, similarly, what term we *should use*, rather than a dispute about descriptive issue about what term we are using.

The main difference between the above two characters of a terminological disagreement is the following. Whereas a *terminological dispute* captures merely the difference between *descriptive* claims (e.g. what term *is* chosen) as made by the parties to the dispute, a *terminological negotiation* captures the difference between *normative* claims (e.g. what term *should be* chosen) as made by the parties to the dispute.

It is worth noticing that different *factors* such as semantics, pragmatics, compositionally, lexical effects, etc., connected to the term can play a part in a terminological disagreement. Moreover, these factors may have different weight in different *contexts*, e.g. theoretical vs. practical vs. legal. For example, consider the following far–reaching consequences of the legal wars over *trademarking* (or *copyright*) of certain terms. When what is called 'Greek yogurt' took over Europe and the United States, many people in Greece had never heard of the term 'Greek yoghurt.' Instead, they have been using the term '*straggisto*' (strained yogurt). At that point, the yogurt in question had never been patented by the Greek state or any Greek company (unlike 'feta cheese' which is now a protected EU term). Eventually, the Turkish company Chobani decided to label their yogurt 'Greek.' In the next seven years, their Greek

---

[99] "A dispute over S is (broadly) verbal when, for some expression T in S, the parties disagree about the meaning of T, and the dispute over S arises wholly in virtue of this disagreement regarding T." (Chalmers 2011, 522)

[100] "A metalinguistic negotiation is a metalinguistic dispute that concerns a normative issue about what a word should mean, or, similarly, about how it should be used, rather than the descriptive issue about what it does mean." (Plunkett 2015, 828)

yoghurt has become the best selling yogurt in the United States. However, the Greek company Fage that produces most of the yogurt that resembles the Greek strained yogurt took the Turkish company Chobani to court for *falsely labelling* its products as Greek. Interestingly enough, Chobani conceded defeat in the United States but won the case on appeal in the United Kingdom in January 2014. In result, the Greek company Fage now labels its product as 'authentic Greek yogurt' in the United Kingdom whereas Chobani labels its yogurt as 'strained' in the United States but can still call it 'Greek' in the United Kingdom. Similarly, in order to avoid and circumvent potentially similar restrictions, another non-Greek company, namely a Norwegian company Synnøve that also calls its yogurt 'Greek Yoghurt' (Norwegian: 'Gresk Yoghurt') decided to produce their Greek yoghurt in Greece.

There is a lot more to be said about terminological disagreement, its resolution and its relevance in scientific, academic, legal and everyday discourse. However, offering a general theory for resolving terminological disagreement goes beyond the scope of this chapter. Instead, in this chapter, I work towards a terminological dispute resolution by arguing for the *theoretical term* that *should* be chosen for the philosophical method behind the Unity Assumption (UA). Nevertheless, I hope that at least some of the issues and lessons will generalize to a number of case studies from other areas, too.

In the rest of this chapter, I evaluate two prominent terminological choices, i.e. 'conceptual engineering' (see Scharp 2013; Eklund 2015; Cappelen 2018; Burgess, Cappelen and Plunkett 2020; Chalmers 2020) and 'conceptual ethics' (see Burgess and Plunkett 2013a,b; Plunkett and Sundell 2013; Burgess and Plunkett 2020), against the two main conditions that I suggest one should take on board when making terminological choices for the philosophical method behind UA: (i) *the semantic correctness condition* drawing on Belleri (2018), and (ii) *the beneficial lexical effects condition* drawing on Cappelen (2018) and Chalmers (2011, 2020). Finally, I engage in a *terminological dispute resolution* by arguing that we should introduce a new term, 'philosophical engineering.' For the philosophical method in question this terminological choice, I argue, fares better against its competition with respect to both of the above mentioned conditions.

In the next section, I elucidate and narrow down the terminological disagreement I focus on in this chapter from other terminological disagreements in the vicinity.

3.2 Terminological Candidates

Theorizing connected to UA goes by different names such as 'conceptual engineering' (see Blackburn 1999; Brandom 2001; Scharp 2013; Eklund 2014, 2015; Cappelen 2018; Isaac 2020; Chalmers 2020), 'conceptual ethics' (see Burgess and Plunkett 2013a,b; Plunkett and Sundell 2013; Burgess and Plunkett 2020), 'verbal disputes' Chalmers (2011), 'ameliorative projects,' 'analytical projects' (see Haslanger 1999, 2000, 2006), 'revisionary projects' (see Railton 1989, 1993; Scharp 2007, 2013),

'explication' (Carnap 1947; Quine 1951; Brun 2016, 2017; Dutilh Novaes 2018; Thomasson 2020), 'gradual semantic drift' (Dorr and Hawthorne 2014); 'contextual and meaning negotiation' (Ludlow 2014); 'metalinguistic negotiations' (Plunkett and Sundell 2013; Plunkett 2015), etc. The above mentioned terminology is often used interchangeably for different theoretical purposes and levels of theorizing including *domain-specific*[101] *theorizing*, *general*[102] *theorizing*, or *theorizing about the philosophical method*[103] *behind UA*.

It might be good not to group all of the different labels associated with levels of theorizing connected to UA together, since although some of them might be grouped together, there are significant differences between others.[104] However, my intention here is to indicate that the substantive issues behind the above labels that parties to the dispute are using can sometimes be seen as competitors, sometimes as compatible, and sometimes as talking past each other. Even though an independent point can be made about the ambiguous and overlapping use of the above mentioned terminologies for different theoretical purposes, in this chapter, I am interested in the recent attempts to resolve a terminological disagreement[105] about the *philosophical method behind UA*. My concern in this chapter is to offer a terminological resolution of the dispute about the *philosophical method* (and the field) that is common to those who engage in FOP, in contrast to those who use these labels at different levels of theorizing, e.g. for *domain-specific theorizing*, *general theorizing* as opposed to *meta-theorizing* (see Chapter 1). Due to the limited scope of this chapter, I look at the two[106] most prominent terminological candidates that have recently been used for the philosophical method behind UA: 'conceptual engineering' and 'conceptual ethics.'[107]

---

[101] See Haslanger (2012) on GENDER and RACE; Scharp (2013) on TRUTH.

[102] See Carnap 1947; Haslanger 1999, 2000, 2006; Scharp 2007, 2013; Chalmers 2011; Dorr and Hawthorne 2014; Eklund 2014, 2015; Ludlow 2014; Plunkett 2015; Brun 2016, 2017; Cappelen 2018; Dutilh Novaes 2018; Thomasson 2020.

[103] See Cappelen 2018; Burges, Cappelen and Plunkett 2020; Burgess and Plunkett 2020; Chalmers 2020; Isaac 2020; Scharp 2020.

[104] Thanks to Sarah Sawyer for pointing this out.

[105] I leave the exact character of these disagreements aside. However, my proposed resolution of a terminological disagreement about UA has a normative character, namely I take myself to participate in a terminological negotiation (see Section 5).

[106] Here is a brief historical note on some of the early uses of the two terms. The coining of term 'conceptual engineering' should be credited to Richard Creath (1990). The term has been popularized by Blackburn (1999, 2). However, the origins of the term can, at least partly, be traced back to Carnap and his use of 'linguistic engineering.' It has been carried on mainly within Carnapian scholarship and is connected to the more contemporary usage of the term by e.g. Brandom, Dretske, Fodor and Milikan (see Isaac 2020, 4). On the other hand, the term 'conceptual ethics' has a shorter history. It was coined by Burgess and Plunkett (2013a) and it has been endorsed further in e.g. Burgess and Plunkett (2013b); Burgess (2014); Plunkett and Sundell (2013); Burgess and Plunkett (2020).

[107] It is, however, worth noting that following Burgess and Plunkett's (2020, 6) latest definition of the relationship between conceptual ethics and conceptual engineering, one may not take the term 'conceptual ethics' to stand for philosophical method behind UA. In particular, they give a definition of conceptual ethics as one of the mereological activities of conceptual engineering: "conceptual engineering is indeed definable in terms of these three activities (conceptual ethics, conceptual innovation, and conceptual implementation)" (Burgess and Plunkett 2020, 6).

In particular, I focus mainly on an explicit terminological disagreement about the philosophical method behind UA that can, most recently, be found in Cappelen and Plunkett's introduction to the volume titled *Conceptual Ethics and Conceptual Engineering* (Burgess, Cappelen, Plunkett 2020) as well as in some of their relevant earlier work. In the above mentioned introduction, Cappelen and Plunkett (2020) offer the following motivation and justification for why they allow the contributors to the above mentioned volume to use the terminology in different ways:

> 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' and 'conceptual ethics', and '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.
> (…)
> *The title of this volume uses two expressions to describe its topic:* 'conceptual engineering' and 'conceptual ethics'. Why? The answer to this is not straightforward. 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. (Cappelen and Plunkett 2020, 2; italics mine)

However, despite endorsing a *terminological pluralism* to be used among the contributors of the above mentioned volume, in the rest of the introduction, Cappelen and Plunkett (2020) engage in what I identify as a *terminological negotiation* about the philosophical method behind UA. To resolve their terminological disagreement they propose a *two-fold* solution. As they put it: "(Cappelen) likes to use the expression 'conceptual engineering,' whereas the other one (Plunkett) thinks that a number of the issues involved are best described as issues in 'conceptual ethics' rather than 'conceptual engineering'" (Cappelen and Plunkett 2020, 2–3).

Before proceeding further, it is worth noting that even though Cappelen and Plunkett (2020) may engage not only in a *terminological* but also in a *merely verbal* and *substantive disagreement* about the philosophical method behind UA, in the rest of this chapter, I focus primarily on the details of their *terminological disagreement*. In the next section, I evaluate the two prominent terminological choices against the *semantic correctness condition*, the first of the two main conditions which, I believe, one should take into consideration when offering a terminological resolution about the philosophical method behind UA.

3.3 Semantic Correctness Condition

Belleri (2018) utilizes the following definition of a merely verbal dispute (MVD+) she borrows from Jenkins (2014): "(MVD+) A dispute is merely verbal iff: (i) the parties are engaged in a prima facie genuine dispute D on a certain subject matter S; (ii) the parties do not disagree on S; (iii) they appear to

disagree on S because of divergent uses of language" (Jenkins 2014, 21). Furthermore, Belleri (2018) distinguishes between two different species of the same genus of merely verbal disputes: *faultless* and *faulty*. Below are some preliminaries of her account.

First, her distinction between *faultless* and *faulty* depends on how one answers the *semantic correctness question*: "[Semantic correctness question] Are the linguistic uses each party is making semantically correct (in some relevant language L)?" (Belleri 2018, 695)

Second, she defines the *semantic correctness* of the use of words (or sub-sentential expressions) in the following way: "[Semantic correctness] Use of an expression is semantically correct just in case such use conforms with what, on the basis of usage (of a certain language L), are considered the expression's meaning and content" (Belleri 2018, 696).

Third, Belleri understands *meaning* "along the lines of a dictionary entry: a suitably stable and general set of conditions (which do not have to be necessary and sufficient) that a competent speaker can (but need not) cognitively entertain" (Belleri 2018, 696) and she understands *content* "along the lines of an extension, that is, an individual or a set of individuals identified in terms of the conditions spelled out by the meaning; this constitutes the world-related component of an expression's semantic[108] profile" (Belleri 2018, 696). To illustrate her distinctions, Belleri (2018, 696) offers the following example claiming that:

> … when faced with an utterance of "Stars are shiny," we shall say that use of "star" is semantically correct just in case it conforms with what, on the basis of usage, we identify as the meaning of "star" (say: "energy-irradiating mass of gas located at a remote distance from Earth"), as well as with its content or extension (the celestial objects that actually fall under this definition). (Belleri 2018, 696)

Lastly, a *faultless*[109] *merely verbal dispute* is one which satisfies the conditions of a merely verbal dispute (MDV+) and also gives a positive answer to the semantic correctness question (see Belleri 2018,

[108] Belleri (2018) focuses primarily on literal uses. However, she believes that her distinction could be applied to non-literal uses as well: "Considering non-literal uses as well would probably require pragmatic correctness—definable as a use's successful communication of a content that is reasonable given the context of utterance. Even with pragmatic correctness, we could still distinguish between merely verbal disputes that are faultless (because both uses are pragmatically correct) and merely verbal disputes that are faulty (because at least one is pragmatically incorrect)" (Belleri (2018, 696).

[109] Belleri (2018) offers the following example: "Mary, speaking British English, utters "Footballs are round"; Jerry, speaking American English, replies "Footballs are not round" (the example draws on Manley 2009). As a result of their uses of different idiolects of English, "football" refers to different items in their respective utterances, and the disagreement is merely verbal by the lights of (MDV+); yet both uses are faultless, in the sense that they comply with what "football" means in British and American English, respectively" (Belleri 2018, 697–698).

697–698). A *faulty*[110] *merely verbal dispute* is one which satisfies the conditions of a merely verbal dispute (MDV+) while giving a negative answer to the semantic correctness question (see Belleri 2018, 697-698).

Given the type of a terminological disagreement that is the focus of this chapter, it is worth mentioning that Belleri's (2018) interest in these two species of merely verbal disputes is not primarily for those disputes that occur in everyday situations but rather for those that occur in a *theoretical* setting, such as *philosophical disputes* (see Belleri 2018, 697–699). She, for instance, utilizes this distinction for the purposes of analyzing a philosophical dispute between universalism and nihilism in metaontology.

I want to bring into connection our terminological choices and merely verbal disputes. As a starting point, take an observation that the speakers are not ideal reasoners and that terms can affect us in different ways. As Chalmers puts it: "Ideal agents might be unaffected by which terms are used for which concepts, but for non-ideal agents such as ourselves, *the accepted meaning for a key term will make a difference to which concepts are highlighted, which questions can easily be raised, and which associations and inferences are naturally made*" (Chalmers 2011, 542). Given the above observation and the close connection between terms, concepts and the standard linguistic meaning, one may argue that one of the major factors that contributes to a development of either faultless or faulty merely verbal dispute are our terminological choices. For example, the *term* (e.g. 'star') we choose for a certain *concept* (e.g. STAR) can comply (or fail to comply) with a certain standard linguistic meaning and, thus, contribute to a certain verbal dispute being faultless or faulty.

In this respect, I suggest to further distinguish between a *faultless* and *faulty terminological dispute*. A *faultless terminological dispute* occurs when the *term* we choose for the *concept* complies with the *standard linguistic meaning*, which, in result, leads to a *faultless verbal dispute*. For example, choosing the term 'star' (in English) and the term 'zvijezda' (in Croatian) for the same concept with the same standard linguistic meaning are *faultless* terminological choices in English and Croatian language, respectively. *A faulty terminological dispute* occurs when the term we choose for the concept does not comply with the standard linguistic meaning, which, in result, leads to a faulty verbal dispute. For example, choosing the term 'star' for the concept MOON does not share the standard linguistic meaning connected to the term 'star' and is, thus, a *faulty* terminological choice in English language.

Understanding these two different varieties of terminological dispute resolution is important because, just like a faulty verbal dispute resolution, a faulty resolution of a terminological dispute can be *harmful*. A *faulty terminological dispute resolution* can, for instance, be an impediment to understanding

---

[110] Belleri (2018) offers the following example: "To illustrate, imagine that Jerry, a native English speaker, is using "star" to mean any celestial body visible from Earth with the naked eye—with the exception of the moon and occasional comets (the example draws on Szabó 2008). He utters, "Venus is a star." Mary is using "star" with its current meaning and utters, "Venus is not a star." Jerry's use is faulty, since it is not warranted by current theories of what a star is, which inform the standing meaning of "star" in current English. The dispute is merely verbal by the lights of (MVD+), for the prima facie disagreement about Venus does not obtain, given how Jerry uses the word "star." Plus, it is a faulty merely verbal dispute, since his use fails to comply with the standard linguistic meaning of "star"" (Belleri 2018, 698).

(see Chalmers 2011, 517), an impediment to communication (see Cappelen 2018, 132–133), a cognitive impediment (see Isaac 2020, manuscript), or it can have legal consequences that may lead to certain practical inconveniences (consider the above example about 'Greek Yoghurt').

In the rest of this section, I evaluate how the terminological choices 'conceptual engineering' and 'conceptual ethics' of the parties to the prominent terminological dispute about the philosophical method behind UA fare with respect to *faultlessness* and *faultiness*.

'Conceptual Engineering'

In *Fixing Language: An Essay on Conceptual Engineering*, Cappelen (2018) advocates for the term 'conceptual engineering.' His use of the term 'conceptual engineering' is *doubly non-descriptive*. As he admits: "I've settled on 'conceptual engineering' though it is far from ideal. It's important that readers *don't take that name as a description*: on the view I defend in this book, the project isn't about concepts and there isn't really any engineering" (Cappelen 2018, 4; italics mine). It is, however, important to stress that there are two ways to interpret Cappelen's (2018) terminological choice since he uses the term univocally both for the philosophical method behind UA as well as for his own general account he dubs the 'Austerity Framework.'

First, when it comes to describing the philosophical method behind UA, for Cappelen (2018), the term 'conceptual' has a non-technical use and can be seen as referring non-descriptively to different subject matters such as concepts, conceptions, topics, terms, words, ideas, etc. However, according to his Austerity Framework, the term 'conceptual' does not stand for concepts. It, instead, stands for (meta)semantic properties such as *intension* (see Cappelen 2018, 50–51).

Second, when it comes to the 'engineering' bit of his terminology, Cappelen (2018) does not believe that there is any kind of engineering going on in the real sense of the word. When he talks about the philosophical method behind UA, the term 'engineering' can be seen as standing for different activities such as revision, replacement, ethics, amelioration, explication, etc. However, according to his Austerity Framework, the term 'engineering' is a kind of activity that causes *reference change* (see Cappelen 2018, 51).

More recently, in the introduction to *Conceptual Engineering and Conceptual Ethics* (Cappelen and Plunkett 2020), Cappelen states that:

> 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. (Cappelen and Plunkett 2020, 3)

Cappelen (see Cappelen 2018; Cappelen and Plunkett 2020), thus, offers a detailed explanation of his terminological choice which reveals that neither part of the terminology 'conceptual engineering' answers the *semantic correctness question* positively. In other words, given Cappelen's own standards, we can conclude that 'conceptual engineering' is descriptively inadequate and that settling terminological negotiation on the term 'conceptual engineering' would be *faulty*.

'Conceptual Ethics'

Burgess and Plunkett (Burgess and Plunkett 2013a; Cappelen and Plunkett 2020; Burgess and Plunkett 2020) advocate for the term 'conceptual ethics' for the philosophical method behind UA. In short, they care about what they call the ethics concerning "what concepts to use—rather than ethics done in an especially conceptual way (whatever that might look like)" (Burgess and Plunkett 2013a, 1098).

When it comes to the 'conceptual' bit of their terminology, they claim that they considered using the term 'representation.' However, the term 'concepts' prevailed because:

> … the fact remains that we do plenty of things with concepts besides using them to describe or represent the world: we ask questions, make plans and promises, bullshit each other, etc. These activities might well be parasitic on the representational functions of our conceptual schemes, but it hardly follows that the norms governing the former can just be read off those for the latter. (Burgess and Plunkett 2013a, 1094)

Nevertheless, Plunkett (Cappelen and Plunkett 2020) recognizes that as "with conceptual engineering, parallel questions can of course arise for other representational devices beyond concepts (e.g., words, if there are any)" (Cappelen and Plunkett 2020, 4).

When it comes to the 'ethics' bit of their terminology, both in Burgess and Plunkett (2013a) and in Cappelen and Plunkett (2020), the choice of the term 'ethics' is supposed to cover a broader conception of ethics than usually conceived. As Burgess and Plunkett (2013a) put it, it covers "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." In fact, even though they do not equate ethics and moral philosophy, they believe that, for the conception behind the term 'ethics' they use, a "closer approximation might be 'practical' philosophy" (Burgess and Plunkett 2013a, 1094).

However, Plunkett (Cappelen and Plunkett 2020) in the introduction to 'Conceptual Engineering and Conceptual Ethics' admits that the label they use is not 'ideal':

> Many will still hear 'ethics' more narrowly—e.g., as tied to distinctively practical norms of the sort that have their home in moral and political philosophy. But other possible labels have their own drawbacks. For example: 'conceptual assessment', another possible label here, makes it sound as this area solely concerned evaluative

> claims about concepts (e.g., which concepts are better than others), leaving out normative claims (e.g., about which concepts an agent should use). (Cappelen and Plunkett 2020, 4–5)

Plunkett and Burgess (Burgess and Plunkett 2013a; Cappelen and Plunkett 2020), thus, offer a detailed explanation of their terminological choice which reveals that neither part of the terminology 'conceptual ethics' answers the *semantic correctness question* positively. In other words, given Burgess and Plunkett's own standards, we can conclude that 'conceptual ethics' is descriptively inadequate and that settling terminological negotiation on the term 'conceptual ethics' would be *faulty*.

To conclude, Cappelen as well as Plunkett and Burgess offer a *faulty terminological dispute resolution* because the terminology they chose for the concept does not comply with the standard linguistic meaning connected to the terms they use which, in result, leads to a faulty verbal dispute resolution. Cappelen's terminological choice 'conceptual engineering' and Plunkett and Burgess' terminological choice 'conceptual ethics' is *faulty* given their own criteria about what the philosophical method behind UA *is* (or *should be*); independently of what the philosophical method behind UA, in fact, *is* (or *should be*).

However, the *faultiness* of a *terminological* and *merely verbal* dispute resolution does not automatically make the relevant *substantive* dispute faulty. As Belleri (2018, 704) points out: "although some of the involved uses may count as semantically incorrect given current usage of the English language, it might be philosophically worthwhile to engage in a metalinguistic argument aimed at establishing whether or not we should change the current usage" (Belleri 2018, 704). Nevertheless, for the present purposes, it is enough to show the faultiness of their terminological resolution given that a faulty terminological and merely verbal dispute resolution can be harmful.

Finally, a terminological dispute resolution can be harmful not only when it leads to a faulty terminological and verbal dispute resolution (possibly resulting in the above mentioned impediments) but also when it leads to *harmful lexical effects*. In the next section, I evaluate the two prominent terminological choices against the *beneficial lexical effects condition*, the second of the two main conditions which, I believe, one should take into consideration when offering a terminological resolution about the philosophical method behind UA.

3.4 Beneficial Lexical Effects Condition

Drawing mainly on Davidson (1978) and Lakoff (2004) on metaphor, Cappelen (2018) offers the following definition of a *lexical effect*:

> Lexical effect—i.e., an effect that words have that go very far beyond anything they say, implicate, or presuppose.

> That is, words have effects that go far beyond their semantic and pragmatic properties (unless you like to use the word 'pragmatic' to denote non-semantic effects, which is not how I use it in this book). (Cappelen 2018, 126)

Cappelen (2018, 123–124) illustrates lexical effects by considering linguistic expressions such as *personal names*, *pejoratives*, and *brand names*. For example, he invites us to imagine today's parents naming their newborn 'Hitler' without being aware of the evil bearer of that name. Next, he invites us to imagine a case where there is "no semantic, pragmatic, or other connection between the child and the evil bearer of the same name." However, he claims, "the name will doubtlessly have emotional and cognitive effects on those who are aware of its former bearer" (Cappelen 2018, 123–124).

Furthermore, Cappelen stresses the potential significance of the study of lexical effects as well as that it is deeply empirical (see Cappelen 2018, 130). He offers some conjectures about lexical effects such as a conjecture that there will be both cognitive and non-cognitive kinds of lexical effects, as well as conjectures about their instability across large populations, temporal instability, non-compositionally, variation of their effects across different individuals (see Cappelen 2018, 130–131).

In a nutshell, given that we are non-ideal reasoners, certain terms or lexical items can trigger different sorts of effects that do not always need to be truth-conditional.[111] Moreover, depending on the context of evaluation, one might find certain lexical effects beneficial whereas another pernicious. Given that, it is plausible to imagine that certain terms might have lexical effects that one might, for e.g. practical reasons, want to preserve or avoid. Being aware of this possibility can put one in a position to make better terminological choices. For example, consider a short history of a terminological negotiation that resulted in the term 'power nap.' Originally, the term was coined for the purposes of instituting a policy of taking naps early during long-haul flights among pilots. This policy was based on the biology-based prediction that "inserting a nap at the front end of an incoming bout of sleep deprivation, you could insert a buffer, albeit temporary and partial, that would protect the brain from suffering catastrophic lapses in concentration" (see Walker 2017, 150). The initially proposed term was 'prophylactic nap.' However, despite accepting the above prediction, the US Federal Aviation Authority (FAA) believed that:

> … the term "prophylactic" was ripe for many a snide joke among pilots. Dinges suggested the alternative of "planned napping." The FAA didn't like this, either, feeling it to be too "management-like." Their suggestion was "power napping," which they believed was more fitting with leadership- or dominance-based job positions, others being CEOs or military executives. (Walker 2017, 150)

Even though one might, initially, think that the lexical effects of our terminological choices do not matter for *theoretical* projects, Cappelen (2018) points out that lexical effects are not used only outside of

[111] For example, Cappelen (2020) argues that lexical items play an important cognitive and theoretical role because they can serve as a 'as marker of topic continuity.'

philosophy but are often also used within 'serious' theorizing[112] and that, in fact, it "might turn out that even for us [philosophers], a significant amount of our work consists in trying to trigger the right kinds of lexical effects" (Cappelen 2018, 129–130). Similarly, Chalmers (2020) draws on Cappelen's notion of lexical effects and endorses the usage of terminological choice 'conceptual engineering' by stressing its practical purpose:

> My metaphilosophical ideology, expressed in my article 'Verbal Disputes', is that you shouldn't get too hung up on words. On the other hand, words matter at least for practical purposes. (…) Conceptual engineering is a hot topic, people have conferences on it. So using that name has what Cappelen calls a lexical effect (p. 122). Using only another name, like 'conceptual coinage', is not going to work in the same way. So pragmatically it makes sense to try and attach this thing you're interested in to this word. That's not just a selfish benefit. (…) At the same time, words have power, and 'conceptual engineering' is a powerful expression. If people had just talked of 'concept repair' or 'concept re-engineering' from the start, I don't think the interest would have been so great. The power comes in part because of the suggestion of something that goes beyond mere fixing. I think we should use the term in a way that vindicates that power. (Chalmers 2020, 8–9)

Interestingly enough, Cappelen (2018) argues against *exploiting* lexical effects. In fact, according to him, exploiting lexical effects should be *banned* (see Cappelen 2018, 133). For, he believes, their exploitation carries the risk of "the possibility of verbal disputes (and fake agreements) and more generally breakdowns and discontinuity in communication" (Cappelen 2018, 132). Moreover, he believes that it is complicated to find the balance between the value of lexical effects and the dangers of miscommunication (see Cappelen 2018, 133). He calls lexical effect-exploiting engineers *exploiters* and claims that they are the 'bad boys' of conceptual engineering (see Cappelen 2018, 149). In the rest of this section, I show that Cappelen's terminological choice 'conceptual engineering' as well as Burgess and Plunkett's terminological choice 'conceptual ethics' about the philosophical method behind UA fall prey to the predicament of being *lexical effect-exploiting engineers*.

'Conceptual Engineering'

Cappelen believes that despite descriptive inadequacy (see the previous section), there are reasons why we should use his terminology for the philosophical method behind UA. In the introduction to *Conceptual Engineering and Conceptual Ethics* he admits: "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). He draws on his previous work (Cappelen 2018) where he claims that 'conceptual engineering' is: "*the label [that] directs people's attention in the right direction* and that's why I've chosen to

[112] "It would be an absurd prejudice to assume that *those engaged in 'serious' theorizing are immune to lexical effects*. They are obviously not. How we label our views and the choice of central theoretical terms can have all kinds of non-cognitive effects that scholars often exploit (for the most part without being aware of doing it). The use of a particular term can be a way to indicate allegiances, trigger associations, make appeal to authority (some famous person used this word), and sometimes just to show off." (Cappelen 2018, 128; italics mine)

go with the Blackburn/Eklund terminology" (Cappelen 2018, 4; italics mine). In more detail, in Cappelen (2018), he considers three specific reasons for using the label 'conceptual engineering':

> So why use the expression 'conceptual engineering' when it turns out to be so misleading? The most important reason is that this label best captures the self-image of the various philosophers that I describe in Part I. (…) in those theories 'concept-like' entities often play a role. (…) it's familiar to the people and the traditions that the book theorizes about. (…) second and related reason: my view helps bring out that many conceptual engineers have been mistaken about what they've been doing, mistaken about how to go about doing what they want to do, and also about the very possibility of doing what they want to do. (…) third reason for using the label 'conceptual engineering' is that I hope this book will encourage others to provide competing frameworks. (Cappelen 2018, 199–200)

Below I consider some immediate drawbacks of Cappelen's terminology that go against his above mentioned reasons.

Against Reason One: Better Self-image. The label 'conceptual engineering' is *misleading* when it comes to capturing the self-image of the traditions that the book theorizes about since examples in the book in many cases go beyond concepts and engineering, including Cappelen's own view. Furthermore, the label, arguably, becomes even more misleading as the existing conceptions of the philosophical method behind UA outgrow concepts as the main objects of engineering.

*Against Reason Two: Realizing One's Mistake*s. This reason operates on the assumption that if one adopts Cappelen's Austerity Framework, one will realize that she has been mistaken about thinking that conceptual engineering is about concepts and engineering. However, I believe, the kind of beneficial lexical effect Cappelen has in mind is *limited*. In particular, the intended beneficial effect does not work on its own. Instead, it might work only for those who are already familiar with the details of Cappelen's view according to which the label is not (only) about concepts and engineering. However, for those who are not on board or are not familiar with his view, the label misses its intended beneficial effect.

Against Reason Three: Competing Frameworks. The idea behind the third Cappelen's reason is that the label 'conceptual' may spark debates and encourage those, who already believe that what we are engineering are concepts, to provide alternative, competitive accounts that go beyond concepts. Even though there has been a significant uptake in the literature when it comes to the term 'conceptual engineering,' I believe that the term 'conceptual engineering' has somewhat limited chances of encouraging people to provide competing frameworks which take objects of engineering other than concepts or concept-like objects since the term 'conceptual' suggests narrowing down the objects of engineering to concepts or concept-like entities. In Chapter 2, Section 2.2.2, I defended the pluralist scope to FOP objects that allows for FOP to be significantly different than concepts and concept-like entities.

'Conceptual Ethics'

When it comes to the term 'conceptual,' Burgess and Plunkett (2013a) explain that they didn't choose the alternative term 'the ethics of representation' because "the R-word seems to carry strong inflationary connotations, and we would like to remain neutral on deflationism about truth and reference" (Burgess and Plunkett 2013a, 1094). Moreover, they claim, even though the activities we use concepts for may well be "parasitic on the representational functions of our conceptual schemes" it is not clear that "the norms governing the former can just *be read off* those for the latter" (Burgees and Plunkett 2013a, 1094; italics mine). Furthermore, they offer two benefits of calling the philosophical method behind UA 'ethics'. First, "'ethics' is less clunky than 'normative and evaluative theorizing'" (Burgess and Plunkett 2013a, 1094). Second, they believe that 'practical' philosophy could have negative lexical effects because "we should keep in mind that paradigmatically 'theoretical' considerations (like cutting nature at the joints) are often front and center in conceptual ethics" (Burgess and Plunkett 2013a, 1094). Finally, they invoke what they take to be positive practical effects of the term 'conceptual ethics:'

> But why bother baptizing it? Any why try to survey such a sparse and scattered field? Our motivating thought is just that the field would remain fallow so long as no one pointed out its potential—or paved the roads between the plots we surveyed in Section 2. Moreover, having a convenient verbal handle for conceptual ethics increases the odds of spotting normative issues about thought and talk implicit or explicit in the next philosophy paper you read. (Burgess and Plunkett 2013a, 1096)

The above quote explicitly suggests that Burgess and Plunkett are considering the beneficial lexical effects they could exploit by choosing to introduce the term 'conceptual ethics' and can, thus, be seen as lexical effect exploiters.

To sum up, the above discussion suggests that Cappelen's terminological choice 'conceptual engineering' as well as Burgess and Plunkett's terminological choice 'conceptual ethics' was made with awareness and intention of exploiting lexical effects of the respective terms. More importantly, their tendencies to exploit lexical effects in combination with their *faulty* resolution of a terminological negotiation about the philosophical method behind UA, puts them in the company of the kind of lexical effect exploiters that Cappelen (2018, 149) calls the 'bad boys' of conceptual engineering that he argues should be banned. In the final section, I introduce a more permissible strategy for lexical effect exploiters that Cappelen (2018) does not consider. Finally, I suggest applying this strategy to offer an alternative resolution of the terminological dispute about the philosophical method behind UA.

3.5 Introducing 'Philosophical Engineering'

A Good Exploiter of Beneficial Lexical Effects

In contrast to the strategy called by Cappelen (2018) the 'bad boys of conceptual engineering' or, to put it gender neutrally, a 'bad exploiter of beneficial lexical effects' (BEBLE), I want to introduce a strategy to be used by a different kind of a lexical effect exploiter, call it a 'good exploiter of beneficial lexical effects' (GEBLE). It is important to stress that *good* here does not mean that it is enough for an exploiter to merely have good intentions while exploiting lexical effects since good intentions *per se* may still lead to miscommunication. As Cappelen (2018) notices:

> There are of course Exploiters with good intentions, but the overall effect of their exploitation is to contribute to and encourage a use of language that undermines what we should treasure the most about it: the continuous exchange of ideas. Exploiters are in effect anti-intellectualist opportunists that contribute to a destruction of genuine communication. (Cappelen 2018, 133–134)

I do agree with Cappelen that, in their aims to achieve beneficial lexical effects, some exploiters of lexical effects with *good intentions* can still be *bad* or BEBLEs. However, I do not think that all exploiters always need to be anti-intellectualist opportunists. Instead, I suggest GEBLE to differ from BEBLE in that she meets the following two constraints: (i) GEBLE aims to comply with the semantic correctness condition; (ii) GEBLE explicitly embarks on a search for beneficial lexical effects.[113] In contrast to BEBLE, GEBLE aims to preserve communication while exploiting beneficial lexical effects. GEBLE aims to look for terms that will stress or highlight crucial parts of what is to be communicated without being misleading or being an impediment to understanding, on pain of achieving the beneficial lexical effects.

One may argue that the GEBLE's strategy is more limited than BEBLE's since the semantic correctness constraint narrows down the *current terminological pool*. This limitation can be serious because, in some cases, the semantics connected to the term as well as the term (lexical item) itself can lead to different (beneficial or pernicious) effects. Consider the following two problematic possibilities. First, a *faultless terminological resolution* does not exclude *pernicious lexical effects*. In other words, some of the currently existing terms that comply with semantic correctness may not lead to beneficial lexical effects. In fact, in some cases it may turn out that they lead to pernicious effects. Consider, for example, the above mentioned Cappelen's (2018) hypothetical example of a baby named 'Hitler' which despite complying with semantic correctness condition can, arguably, lead to pernicious lexical effects. Second, a *faultless terminological resolution* does not exclude *pernicious semantic effects.* This is what Cappelen (2018, 33–34) calls 'objectionable effects of the semantic value.' For example, consider *semantically non-*

[113] What in particular *beneficial lexical effects* are needs to be further specified. It could, for instance, differ from one context to another, as long as they do not contradict with the first, semantic correctness, constraint. Moreover, in some cases, beneficial and pernicious lexical effects may compete. Nevertheless, there are different ways of resolving this potential issue, such as cost-benefit analysis strategy.

deficient concepts (because, e.g. their definition is stipulated) such as MARRIAGE, or RAPE which can, nevertheless, still be deficient in other ways, e.g. morally, socially or politically, when excluding the same sex couples to be married, or not counting unwanted sex within marriage as rape (see Cappelen 2018, 28).

In order to address this worry, I believe that GEBLE has at least two broad strategies at her disposal. First, she may turn to engineering, i.e. changing or revising the concept and/or semantics connected with the existing terminological pool. It is, however, worth noting that a terminological change or revision is different from conceptual and/or semantic change or revision since a change or revision of a concept and/or semantics may not require a change or revision in terminology, and, possibly, *vice versa*. Second, into the existing terminological pool GEBLE might introduce a new term connecting it to a new concept and/or semantics, or introduce a new term connecting it to an existing concept and/or semantics, or introduce a new term that comprises existing terms, where the existing terms keep their existing concepts and/or semantics. As an example, one may consider *appropriation of certain slurs*. Despite compliance with semantic correctness, slurs are considered to have pernicious effects. One may argue that *appropriation of slurs* retains the same lexical item while changing the underlying concept, semantics, pragmatics or lexical effects (depending on one's theory of slurs[114]).

It is also worth noticing that terminological, conceptual, and semantic change can often go hand in hand and that both of the above proposed strategies can be seen as instances of philosophical method behind UA. This is, what Chalmers (2020), at least in part, captures under the terms *heteronymous* (different-word) and *homonymous* (same-word) conceptual engineering. As he puts it:

> Of course projects in heteronymous conceptual engineering could be carried out homonymously, and vice versa. Instead of introducing the new word 'supervenience', we could say, here is what 'reduction' should mean. (…) The same goes vice-versa for various homonymous projects, like the revisionary analysis of truth. Kevin Scharp ends up using different phrases, 'ascending truth' and 'descending truth'. (…) But the practical effects will be different. The upsides and downsides here are mainly practical rather than theoretical, although the lines can blur in some cases. (…) sometimes we can achieve our purposes better with a new word, and sometimes with an old word. (Chalmers 2020, 11)

In the rest of this section, I utilize the GEBLE strategy of introducing a *new* term comprising *existing* terms with their *existing* semantics. I apply this strategy to the case of terminological negotiations about the philosophical method behind UA. In particular, I suggest introduction of a *new*

[114] The pernicious effects are often considered to stem either from their *non-semantic* (in Fregean scholarship) or *semantic* (truth-conditional) (in Dummett's scholarship) features.

term, 'philosophical engineering,'[115] which comprises existing terms with their *existing* standard linguistic meaning. With this in place, what remains to be shown in the rest of this section is that the term 'philosophical engineering' satisfies both *semantic correctness condition* and *beneficial lexical effects condition*.

Satisfying the Semantic Correctness Condition

When it comes to satisfying the *semantic correctness condition*, my terminological choice, 'philosophical engineering,' can be seen as *doubly descriptive*. In this chapter, there is not space to restate my full argument for this choice of terminology. In summary form, my motivations for this choice are as follows. In Chapter 2, I argued that none of the current proposals can settle the substantive dispute about philosophical method behind UA. The reason for that is, I argued, that the so-far proposed theories for the philosophical method behind UA cover *the unifying subject matter* and *the unifying activity scope* only partially. This further suggests that a substantive disagreement could not be resolved by settling on one of the current proposals. Instead, by arguing for a more plausible scope of the unifying subject matter and the unifying activity behind UA, I offered another way to resolve the substantive disputes over the unifying subject matter and unifying activity.

In particular, I argued that we should consider '*being philosophical*' as the unifying subject matter behind UA. I suggested that what makes the method behind UA *philosophical* is that philosophy as its *subject matter* is being engineered via engineering different philosophical objects, given that the unifying subject matter is often articulated in terms of different 'entities' or 'categories' (e.g. concepts, meanings, extensions, intensions, theories, etc.) that are considered to be 'objects' of UA. Furthermore, I argued that we should consider *engineering* as the unifying activity behind engineering projects in philosophy. I took a strong non-metaphorical stance towards the unifying activity behind UA being *engineering*. I suggested that what makes the method behind UA *engineering* is that it consists in a *five-stage* (*identification, evaluation, strategic planning, implementation, re-evaluation*) *recursive activity* which, being a congenial activity to engineering methods outside philosophy, bears a *family resemblance* and is in *continuity with other engineering fields*. In a nutshell, according to my substantive dispute resolution, *philosophical engineering* is engaging in a five-stage recursive activity that operates on different

[115] Here is a brief historical remark on the use of the term 'philosophical engineering.' I have been using the term 'philosophical engineering' for the purposes argued in this dissertation since 2018. In 2020, however, I discovered that the term 'philosophical engineering' has already been in the existing terminological pool for around two decades. The term was coined by Tim Berners-Lee, a person often referred to as 'the inventor of the Web.' He introduced the term for the first time in a public forum in 2003: "We are not analyzing a world, we are building it. We are not experimental philosophers, we are philosophical engineers." (Berners-Lee 2003) Other notable uses of the term descending from Berners-Lee's use are the title of the paper 'Philosophical Engineering' by Nigel Shadbolt (2007) (which is also the only occurrence of the term in the paper) and the edited volume titled *Philosophical Engineering: Toward a Philosophy of the Web*, edited by Harry Halpin and Alexandre Monnin (2014). However, even though there are many interesting and useful parallels between their uses of the term and the one suggested in this chapter, their uses are rather confused, or, at best, narrower than mine. Given that, their use is potentially semantically faulty and with suboptimal lexical effects.

philosophical objects.

Thus, given my substantive dispute resolution, my terminological choice of introducing a new term 'philosophical engineering,' avoids the misnomer threat since both parts of the phrase 'philosophical engineering' answer the *semantic correctness question* positively. Moreover, this makes my terminological choice descriptively adequate and my settling of terminological negotiation on the term 'philosophical engineering' *faultless*.

Satisfying the Beneficial Lexical Effects Condition

I believe that the introduction and use of a new term 'philosophical engineering' for the philosophical method behind UA could have a range of beneficial semantic and lexical effects. For example, it could contribute to avoiding terminological confusion in discussions; do a better job describing our current practices in the field; capture better the general self-image of philosophers that engage in projects that do not consider concepts as only objects of engineering; contribute to the unification of the methodological field behind UA; open a way for further explorations and articulation of the field; allow for a more illuminating relationship between philosophical engineering and philosophical analysis.

In order to support the plausibility of the above mentioned possible beneficial effects of the term 'philosophical engineering,' in the rest of this section, I draw on a brief analogy from the history of philosophy. Consider Frege's and Russell's notions of logical analysis; Moore's semantic (and of the world) analysis; Carnap's quasi-analysis and explication; Wittgenstein's analysis of the ultimate constituents of propositions and the nature of the world itself; Oxfords's linguistic and connective analysis; logical and metaphysical analysis of the Cambridge School of Analysis (see Beaney 2018, Soames 2003). Despite substantive disagreement and a "lack of widely agreed-upon philosophical analyses" (Horvath 2017, 2), the philosophical method that goes by the name 'philosophical analysis' is still a useful umbrella term since it has become essential in our theorizing and crucial in unifying the field of analytic philosophy. As Soames (2005, 144) puts it:

> *Philosophical analysis* is a term of art. At different times in the twentieth century, different authors have used it to mean different things. What is to be analyzed (e.g., words and sentences versus concepts and propositions), what counts as a successful analysis, and what philosophical fruits come from analysis are questions that have been vigorously debated since the dawn of analysis as a self-conscious philosophical approach. (…) Indeed the variety of positions is so great as to make any attempt to extract a common denominator from the multiplicity of views sterile and not illuminating.
> Nevertheless analytic philosophy—with its emphasis on what is called "philosophical analysis"—is a clear and recognizable tradition. (Soames 2005, 144)

To sum up, for the philosophical method in question, the term 'philosophical analysis' has proven

itself as a more adequate term for the general methodological field than the term 'conceptual analysis' since it provides a broader conceptual space and can accommodate different conceptions and subfields of analysis within analytic philosophy. Furthermore, it has played a crucial role in unifying different conceptions and subfields of the field of analysis in philosophy, and has, in that sense, become essential in our theorizing and crucial in unifying the field of analytic philosophy. Analogously, given the structural potential and similarity between the debates about the two philosophical methods, i.e. *analysis* and *engineering*, one may argue that the philosophical method behind UA is in a need of a new umbrella term that will have similar beneficial lexical effects. Given the opportunity, I believe that my terminological choice 'philosophical engineering' can achieve this task better than the above mentioned alternatives.

3.6 Conclusion

This chapter brings to the foreground the phenomenon of a *terminological disagreement*. Drawing on Belleri (2018), I argued that a *faulty* verbal and terminological dispute resolution is *deficient* since it can result in e.g. cognitive, understanding and communicative impediments. In order to prevent these impediments, I argued, our terminological choices should comply with the *semantic correctness* whereas the violation of this compliance should be banned or (at least) strongly avoided, especially in the theoretical setting. On the other hand, I argued, our terminological choices may be compatible with *exploitation of beneficial lexical effects* when not on pain of *semantic correctness*.

As a case study, I evaluated the prominent terminological dispute resolution about the philosophical method behind UA as *deficient*. I have shown that, relative to their respective substantive dispute resolutions about UA, both Cappelen's terminological choice, 'conceptual engineering,' as well as Burgess and Plunkett's terminological choice, 'conceptual ethics,' are *faulty*. Furthermore, I have shown that their terminological choices have been governed by *exploiting lexical effects on pain of semantic correctness*.

The novel terminological dispute resolution of the philosophical method behind UA I offered in this chapter operates on the assumption that *semantic correctness should trump beneficial lexical effects* since terminological choices of exploiters of lexical effects who answer semantic correctness question negatively may lead to various of the above mentioned impediments. I have shown that, relative to my substantive dispute resolution of UA, my terminological choice, 'philosophical engineering,' is *faultless*. By meeting both the *semantic correctness condition* as well as the *beneficial lexical effects* condition, I argued, a terminological choice of introducing a new term 'philosophical engineering' for the philosophical method behind UA fares better against its competition.

CHAPTER 4

OBJECTIONS AND REPLIES[116]

Abstract

In this chapter, I address some objections concerning the substantive and the terminological dispute resolutions of the Unity Assumption (UA) that I offered in Chapter 2 and Chapter 3, respectively. In particular, with regard to the substantive dispute resolution, I discuss: Explanatory Power of Philosophical Engineering; Philosophical Engineering Outside of Philosophy; and Deutsch's Pessimistic Dilemma. With regard to the terminological dispute resolution, I discuss: Semantic Correctness or Descriptive Adequacy; Aptness to Cause Misunderstanding, Miscommunication, or Harm; Too Broad and Too Narrow Terminological Choice; What 'Philosophical' Means is Just as Unclear as What 'Conceptual' Means; and No Decisive Reasons for Introducing 'Philosophical Engineering.'

[116] For their critical remarks and help in formulating the above mentioned worries, I would like to thank Farbod Akhlaghi-Ghaffarokh, James Andow, Derek Ball, Herman Cappelen, Paul-Mikhail Catapang Podosky, David Chalmers, Josh Dever, Manuel Gustavo Isaac, Patrick Greenough, Sam Kang, Steffen Koch, Louis Kohlmann, Jennifer Nado, Mark Pinder, David Plunkett, Sarah Sawyer, Elisabeth Cantalopemass, Kevin Scharp, and Davide Andrea Zappulli.

4.1 Addressing Worries about the Substantive Dispute Resolution

In Chapter 2, "Substantive Dispute Resolution of the Unity Assumption," I offered a resolution of the *substantive* dispute behind UA. I argued that none of the current proposals can settle the substantive dispute about UA. The reason for that is, I argued, that the so-far proposed theories for the philosophical method behind UA cover *the unifying subject matter* and *the unifying activity scope* only partially. This further suggests that a substantive disagreement could not be resolved by settling on one of the current proposals. Instead, by arguing for a more plausible scope of the unifying subject matter and the unifying activity behind UA, I offered another way to resolve the substantive disputes over the unifying subject matter and unifying activity that comprise the unifying methodology of FOPs. I engaged in a *substantive dispute resolution* of UA as follows. First, I argued that we should consider *being philosophical* as the unifying subject matter behind UA. More specifically, I suggested that what makes the method behind UA *philosophical* is that philosophy as its *subject matter* is being engineered. This suggestion relies on the assumption that philosophy as a FOPs' subject matter can be engineered via engineering philosophical objects. Second, I argued that we should consider *engineering* as the unifying activity behind UA. More specifically, I suggested that what makes the method behind UA *engineering* is that it consists in a *five-stage* (*identification, evaluation, strategic planning, implementation, re-evaluation*) *recursive activity*. This suggestion takes engineering activity in a non-metaphorical sense and endorses engineering as a philosophical method that bears a *family resemblance* and is in *continuity with other engineering fields*. In a nutshell, *philosophical engineering* is engaging in a five-stage recursive activity that can operate on different philosophical objects.

In this section, I address three types of objections related to this chapter: Explanatory Power of Philosophical Engineering (Section 4.1.1); Philosophical Engineering Outside of Philosophy (Section 4.1.2); and Deutsch's Pessimistic Dilemma (Section 4.1.3).

4.1.1 Explanatory Power of Philosophical Engineering

In this section, I address two worries related to the explanatory power of my *substantive dispute resolution of the philosophical method behind UA*, the worries related to the substantive dispute resolution of the *unifying subject matter* and of the *unifying activity* behind FOPs.

Subject Matter Overgeneralization and Over-generation Worry

Given the *pluralism about the philosophical objects of engineering* that I advocate in Chapter 2, Section 2.2.2, one may argue that there is a danger of *overgeneralization* of what counts as a philosophical object. For instance, one could claim that it seems that while things such as people, numbers, money, animal rights or AI rights may be good candidates for FOPs, things such as nails and eggs may not. It, thus,

seems like there is a serious principal worry that if one allows that a philosophical object is something else than concepts, almost any kind of entity can count as an object of philosophical engineering and almost any kind of subject matter can be a subject matter of the method behind UA. Relatedly, this could end up in *over-generation* of philosophical objects, which could result in *too many projects* counting as FOPs. In result, both the *overgeneralization* and *over-generation worry* could contribute to lowering down the explanatory power of philosophical engineering.

My immediate response to the *overgeneralization worry* is that the explanatory power of philosophical engineering is worrisome in the above mentioned sense to the extent that what 'being philosophical' *means* or *is* is worrisome. However, this worry is an old worry. It is a worry for philosophy itself, i.e. it is a worry about what objects are philosophical and what is the subject matter of philosophy. In this dissertation, I do not offer a definite answer to what objects are philosophical objects nor what falls under the subject matter of philosophy. Instead, for the purposes of this dissertation, I want to borrow Deutsch's (2020a, 2) positive and negative characterization of *philosophical phenomena* in order to offer one possible answer to what it means to be *philosophical*. Deutsch's positive characterization of a philosophical phenomena is the following: "A phenomenon is a philosophical phenomenon simply by virtue of being of interest to philosophers. Knowledge, freedom, personal identity, etc., are philosophical phenomena because philosophers take an interest in these things; there is no more and no less to being a philosophical phenomenon than that, given how I intend "philosophical phenomenon" to be understood here" (Deutsch 2020a, 2). On the other hand, his negative characterization of philosophical phenomena is the following: "I intend no contrast between philosophical phenomena and, for example, *physical* phenomena. Causation is both a physical phenomenon and a philosophical one, for instance. More generally, some of the many phenomena in which philosophers take an interest are straightforwardly physical phenomena" (Deutsch 2020a, 2). Similarly, for the purposes of this dissertation, I rely on current theorizing in philosophy about FOP objects as data points about what can count as a philosophical object. A tentative taxonomy I provided in Chapter 2, Section 2.2.1, can be used as a starting point. This tentative taxonomy could also be used as a crude inductive argument for the claim that we could expect the taxonomy of philosophical objects to change. After all, the history has shown us that there are no "safe spaces" for what being a philosophical object *is*. One way to look at it is to say that this is up for grabs (or, perhaps, engineering) and may depend on many (often contingent) factors (e.g. see Cappelen (2017) for why philosophers should not do semantics anymore, or Scharp (2020, 409–411) for how subject matter of philosophy has changed and been constricted over time). In result, this means that what is 'philosophical' may be shrinking or expanding as one eliminates or introduces new philosophical objects. For the present purposes, I am, however, staying neutral on what these objects are, as long as there is at least one object

that is philosophical. In other words, although there are currently lots of philosophical objects (because there are lots of things philosophers are interested in), there could e.g. be just one.

Second, I want to briefly address a possible *over-generation worry*, or the worry that my subject matter substantive dispute resolution generates *too many projects* which count as FOPs. Even if, as a result of the substantive dispute resolution that I offer, there are many more (current and potential) projects that count as FOPs, I believe that this worry is not a problem for the philosophical method behind UA. In fact, having as many projects as possible to count as FOPs is a strength rather than a weakness of my proposal. Being categorized as belonging to the same method is not a problematic thing *per se,* as long as FOP candidates are sufficiently similar both with respect to their unifying activity and their unifying subject matter. The more pressing worry, however, is whether categorizing certain projects as belonging to this particular method is illuminating in any interesting and important sense. One way to understand the above claim is to ask if there is any practical use that categorization of certain projects as FOPs that fall under the method of philosophical engineering can be put into. In response, I want to say that the account of philosophical engineering that I offer is an important and illuminating category because it unites and recognizes all projects that are *philosophical* and that are instances of *engineering* (see below). Furthermore, (and somewhat analogously to philosophical analysis) even though the category of philosophical engineering that I endorse offers only a broad understanding of FOPs, it comprises numerous, (domain-)specific accounts as its subcategories, which are more detailed and can provide specific blueprints for each FOP.

The Generality Worry for the Five Stage Recursive Model

One may argue that the *Five Stage Recursive Model* I offer is explanatorily weak because it gives an answer regarding what the unifying activity behind UA is at a very general level and it does not provide any specific way of engineering of a particular philosophical object. In other words, this is a worry about the importance of theorizing about the unifying activity behind FOPs at such a general level.

My response is that, even though the Five Stage Recursive Model is more general than the alternative proposals in the vicinity that focus on engineering of a particular philosophical object (see Chapter 2, Section 2.2.1), it plays at least two important roles. First, it is important for philosophical methodology because it clears the ground and provides the foundation for the general methodological research field as well as its subfields, which offer a more specific account of the same methodology, depending on the object of engineering. In other words, the Five Stage Recursive Model is a model for a general philosophical method and less so a model for any specific application of the method. Its strength lies in that it accounts for the unifying activity that underlies FOPs by addressing the

commonalities in the unifying activity that FOPs share. The Five Stage Recursive Model offers a general foundation for the unifying activity for the method behind UA while staying neutral on certain (meta)philosophical questions. As such, it can incorporate different arguments for different philosophical objects. For instance, given its generality, the Five Stage Recursive Model that I offer is similar to Cappelen's (2020) Master Argument for conceptual engineering. However, Cappelen's Master Argument presupposes certain features of his (2018) Austerity Framework and is in that sense less general and more restrictive than it initially claims to be. Furthermore, one can draw a similar comparison with philosophical analysis whose explanatory power is tied to its usefulness for recognizing projects that belong to the same methodological field of analytic philosophy as well as for informing and giving guidance to its subfields (see Soames 2005). Most crucially, the Five Stage Recursive Model for the unifying activity behind UA is not an alternative account to the more (domain-)specific accounts. Instead, it is a more comprehensive framework that offers preliminary guidance for those who (want to) engage in FOPs. Finally, the importance of the Five Stage Recursive Model, and philosophical engineering more generally, is in its contribution to establishing the place of engineering *within* philosophy alongside other engineerings *outside* of philosophy, as well as pointing at their possible interactions, such as a division of engineering labour (see Chapter 2, Section 2.3.3).

4.1.2 Philosophical Engineering Outside of Philosophy

There are quite a few people who think that FOPs can take place outside of philosophy. For example, as Cappelen (2018) states: "Conceptual engineering is an activity that many people—not just philosophers—engage in" (Cappelen 2018, 5). In other words, it seems that a very large proportion of the cases that conceptual engineers point to as exemplars of the method are not really philosophical, or at the very least are not performed by philosophers. So, for instance, consider the revision of the definition of 'planet.' In general, one could be inclined to count nearly all cases of "tinkering with definitions of technical vocabulary" in various fields as conceptual engineering. So the writers of the DSM manual in psychiatry are engaged in conceptual engineering, for instance. As are legal scholars who tinker with the definition of terms used in legal contexts such as '*mens rea*' or 'collusion,' etc. The method behind UA is something being practiced outside of philosophy as well, and probably more than we (philosophers) are aware of. One could argue that these cases show that 'philosophical' engineering is a misleading label—there's nothing uniquely philosophical about the method. If so, one may argue, not all FOPs are species of the general category *philosophical* engineering.

My reply is that FOPs can take place *inside* of philosophy but that FOPs can also take place *outside* of philosophy, or can (at least partially) be done by non-philosophers, in a similar manner in which projects of philosophical analysis can take place both *inside* and *outside* of philosophy, or projects

belonging to some other disciplines can take place outside of their respective disciplines. Consider the following examples. *Philosophical Analysis.* Certain projects of philosophical analysis that are of primary interest of philosophers, e.g. concepts (and/or properties) such as TRUTH, KNOWLEDGE, or certain moral, social political concepts (and/or properties), are usually preformed by philosophers who draw on the theoretical results from philosophical methodology of philosophical analysis. However, even though philosophers might on average be better equipped to engage in philosophical analysis, philosophical analysis is not reserved only for philosophers since anyone, in principle, can utilize the same methodology. *Applied Ethics.* In e.g. bioethics or in environmental ethics, one can come across different domain-specific projects about how to act good, e.g. how to become a good person or how to be good to other non-human animals. However, these ethical projects are intended to be implemented in different areas such as in private and public life, or in different professions, such as health, technology, law, or leadership. Engaging in these projects can be carried out not only by philosophers but everyone who wants to act well by consulting experts and/or studying the literature that theorizes about these topics (and doing so is often a more time, energy, or quality of choice efficient way to engage in these projects). *Applied mathematic and applying mathematics in everyday life.* The former is the application of mathematical methods by different fields such as physics, engineering, medicine, biology, business, computer science, and industry. The latter is, for example, practicing and applying certain parts of mathematics in school, or calculating how much one needs to pay for their groceries, taxes, etc. In both cases, one is applying the tools of a mathematical method outside of mathematics. *Mechanical engineering.* Consider the relation between a mechanical engineer and the construction workers. The mechanical engineer comes up with a blueprint for the project of engineering. There is a plan for each step that the construction workers will be executing. Construction workers will follow a particular engineering plan and orders from their supervisors in order to implement certain mechanical engineering projects. The above examples, thus, point into different ways in which a certain discipline can *take place outside* of itself.[117]

In a similar manner, engineering philosophical objects is not reserved only for philosophers. Depending on the philosophical object in question, both philosophers as well as non-philosophers can contribute towards philosophical engineering, and *vice versa* (for *division of engineering labour* and *engineering holism*, see Chapter 2, Section 2.3.3). For instance, while philosophers working on philosophical methodology may be responsible for giving a theory of the method of philosophical

[117] It is, however, important to notice that what it means to *take place in* philosophy or in certain other disciplines is open to interpretation, given that philosophy and other disciplines do not have any intrinsic (physical) place where they should occur. Given that, in the case of philosophy, this question may boil down to how one answers questions like whether the project needs to originate from or be of a special interest for philosophers, whether a particular method needs to be preformed by philosophers, or whether performing a philosophical method makes one a philosopher. Moreover, this may open a further question about what comes first, a philosopher or a philosophical method.

engineering, practicing philosophical engineering, i.e. engaging in various stages of specific FOPs (such as coming up with strategies or its implementation) can and in certain cases must be done and further informed by non-philosophers. In this sense, philosophical engineering can take place within philosophical methodology. Furthermore, conceptual engineering can be subject to engineering itself (see Cappelen 2018; Isaac 2020). Furthermore, it can happen in different subareas of philosophy, as well as in human inquiry more generally, in public and private debates, or wherever one engages in FOPs of engineering philosophical objects. For example, the non-philosophers can carry out certain stages of philosophical engineering by following explicitly or deferring to the methodology and tools of philosophical engineering. For instance, a non-philosopher who is working out why a certain concept, e.g. MARRIAGE, is defective and in what way, and then trying to fix it, can be using philosophical tools and implementing the theoretical results of philosophy for the purpose of e.g. social justice. Furthermore, she can be using philosophical tools outside of philosophy for not strictly speaking philosophical purposes. In result, apart from focusing on achieving social justice, she may also be contributing to engineering of a particular object that is also of philosophical interest, i.e. the concept of MARRIAGE. Similarly, the non-philosophers could carry out certain stages of philosophical engineering by deferring to the methodology and tools of philosophical engineering. For instance, activists may be changing the meaning of the concept of MARRIAGE without knowing or needing to know all of the details or doing strategic planning or evaluation themselves (for more examples, see Chapter 1, Section 1.1).

Finally, it is important to notice that so far, most attention has be devoted to the 'philosophical' bit of my terminology (which can be seen as being about the unifying subject matter behind UA, see Chapter 2, Section 2.2). However, a similar worry can be applied to the 'engineering' bit of the terminology (which can be seen as being about the *unifying activity* behind UA, see Chapter 2, Section 2.3). My proposal provides an explanation for why the method behind UA is a philosophical method (on pair with philosophical analysis) as well as explains the relation between philosophical and other kinds of engineerings.

4.1.3 Deutsch's Pessimistic Dilemma

Deutsch (2020a) offers a rather pessimistic dilemma for the method behind UA. He argues that those advocating for the method behind UA are either "ignorant of how conceptual engineering can be implemented, or else it is trivial to implement but of very little value, representing no new or especially fruitful method of philosophizing" (Deutsch 2020a, 1). In this section, after discussing the main challenges stemming from Deutsch's pessimistic dilemma, I present the *hard problem of implementation* and show how my model squares with respect to this problem.

The Dilemma

The first horn of Deutsch's (2020a) dilemma considers the method behind UA to be about changing meaning and reference. Moreover, the first horn of Deutsch's dilemma is a dilemma in its own which includes two main ingredients: stipulation and Grice's (1989) distinction between *speaker-meaning* and *semantic meaning*. In a nutshell, the first horn this sub-dilemma is a claim that if the method behind UA is about *changing semantic meaning*, then it faces a *severe implementation problem* since a mere semantic meaning stipulation will not suffice for semantic meaning change. The second horn of this sub-dilemma claims that in case the method behind UA is about *changing speaker-meaning*, then we do avoid the implementation feasibility problem but utilizing stipulation as a tool to change speaker-meaning makes the implementation trivial and *of very little value*.

The second horn of Deutsch's dilemma considers the method behind UA to be about *stipulative introductions* or *stipulative additions*. Deutsch admits that, in these cases, implementation feasibility is trivial.[118] However, he argues that such a construal of the method behind UA faces another type of problem: even though philosophy as well as other disciplines can benefit from the above kinds of stipulations, this practice is *neither* a *new* nor *neglected method*, both in philosophy as well as beyond.

A couple of remarks about Deutsch's dilemma are in order. First, Deutsch's dilemma concerns five candidate strategies for the main unifying activity of the method behind UA. In particular, the first horn considers: (i) *stipulative revision of semantic reference and meaning*, and (ii) *stipulative revision of speaker's reference and meaning*; whereas the second horn considers: (iii) *stipulative introduction of meaning and reference for a new term*; (iv) *stipulative addition of meaning and reference for an old term;* (v) *revealing or characterizing the meaning or reference of an old term.*[119]

Second, Deutsch evaluates the above mentioned five candidate strategies for the method behind UA against three criteria: (i) *implementation feasibility*, (ii) *implementation value*, and (iii) *methodological newness or negligence*. In particular, Deutsch raises the implementation feasibility challenge for the first candidate strategy whereas he argues that the implementation of the other four candidate strategies to be trivial, namely he sees very little value for the second candidate strategy, and argues that the third, the fourth and the fifth candidate strategies are neither new nor neglected practices.

Third, it is also important to note that out of the five candidate strategies, Deutsch only considers the first candidate strategy, namely the *stipulative revision of semantic reference and meaning*, as a serious

[118] However, Deutsch also admits that: "What is not trivial is the activity of identifying and recognizing phenomena not yet identified or recognized as important to sophisticated philosophical theorizing." (Deutsch 2020a, 12)

[119] See Deutsch 2020a, 10–16.

candidate for the method behind UA.[120] On the other hand, it is worth noticing that my framework allows for all of the above mentioned five candidate strategies as legitimate strategies[121] of the method behind UA (see Section 2.3.2.3). This calls for their defense against Deutsch's challenges.[122] However, I do not take Deutsch's triviality of implementation challenge which says that *implementing strategies (iii)-(v) are neither* a *new nor neglected practice* as a serious objection to my framework. The fact that those strategies might have already been practiced does not speak against my framework as long as they justifiably fit into the framework. In fact, my framework, which takes into consideration the most plausible unifying subject matter and the unifying activity, suggests that these strategies deserve their place within the method behind UA. Recall, the aim of the first part of this dissertation is to work towards *revealing* the nature of the method behind UA, and not to argue that it has not been practiced before. On the other hand, I do take the *lack of implementation feasibility for stipulative revision of semantic reference and meaning* and *very little value of implementation connected to stipulative revision of speaker's reference and meaning* objections more seriously. For this reason, first, I address the *implementation feasibility challenge for stipulative revision of semantic reference and meaning* by showing its prospects within externalist and internalist metasemantics.[123] After that, I turn to addressing the *very little value of implementation challenge connected to stipulative revision of speaker's reference and meaning*. Finally, I address the *hard problem of implementation* for my framework.

The Feasibility Challenge for the Externalist about Semantic Meaning

At first glance, radical metasemantic externalism seems incompatible with stipulative revision of semantic reference and meaning. For example, consider Cappelen's (2018) Austerity Framework which endorses meanings and intensions as primary philosophical objects. His framework is based on a mix of features from the works of Kripke, Putnam, Burge, and Williamson. In particular, Cappelen argues that "while reference change is possible, there is no algorithm for how it is done, meaning that the conditions for successful revision might themselves be in flux" (Cappelen 2018, 72). Furthermore,

---

[120] "But the belief that conceptual engineering might be a fruitful replacement for, or supplement to, a more traditional method, like conceptual analysis, is unjustified, if I am right that conceptual engineering is either (attempted) stipulative revision, or else stipulative introduction or addition. Stipulative revision faces an implementation problem that should make us doubtful of its prospects. And stipulative introductions and additions are common practice already; indeed they are already a *component* of many conceptual analyses. So, in neither case should dressing these practices up by describing them as "conceptual engineering" convince us that they represent a special antidote to what is allegedly ailing more traditional methods of philosophy." (Deutsch 2020a, 15)

[121] As well as some others not mentioned by Deutsch (2020a).

[122] For another angle of dissolving Deutsch's dilemma see Koch 2020b.

[123] The issue, according to Deutsch (2020a), applies both to an externalist and an internalist: "The root of this problem—the implementation problem for conceptual engineering conceived as stipulative revision—is not externalism. On any metasemantic view that requires more, for an existing term to have a particular semantic meaning and reference, than just the intention on the part of some group of speakers to use the relevant term as if it had that very semantic meaning and reference, stipulative revisionists will face the implementation problem" (Deutsch 2020a, 19).

Cappelen promotes a principle he calls "Inscrutable—Lack of Control—Will Keep Trying: The processes involved in conceptual engineering are for the most part inscrutable, and we lack control of them, but nonetheless we will and should keep trying" (Cappelen 2018, 72).[124]

In response[125] to the externalist challenge such as the one posed by Cappelen (2018), Koch (2018) argues that we can achieve *semantic meaning change* by possessing a *long-range* (as opposed to the *immediate) collective* (as opposed *individual) control* over semantic facts. Here is a more detailed picture of how Koch (2018) believes this change should work: "Many people start using the term in question *as if it had the new reference;* eventually, this will add pieces to the body of information we associate with the term that have the new object or kind as their causal source. Likewise, some of the information caused by the previous object or kind will be lost from the body of information. Thus, little by little, the term will shift from the old reference to the new one, until the new object or kind has gained overall dominance" (Koch 2018, 17).

However, Deutsch (2020a) poses a new problem for Koch's (2018) account by arguing that we cannot have a guarantee or a prior knowledge that our collective, long-range efforts to change the way how speakers use a certain term *t*, namely to use it as if it refers to *x*, will turn *x* into the dominant causal source. The reason for that, Deutsch believes, is that speakers can make widespread mistakes, i.e. despite their best intentions to use *t* in relation to *x*, speakers can use *t* in relation to things that are not x (see Deutsch 2020a, 20).

Koch (2020b) dismisses the legitimacy of Deutsch's worry by arguing that even though many of both our individual and collective goals are only within our long-range control and we hardly ever have such guarantees and antecedent knowledge, this usually doesn't stop us from engaging in them, and that this should also be the case with conceptual engineering. As he puts it:

> By contrast, I take it that almost any kind of worthwhile long-term project faces these obstacles (and potentially others), but that this does not and should not stop us from pursuing them. Individuals engage in all sorts of long-term activities without having any guarantee of success or antecedent knowledge about their outcomes. (…) The problem with Deutsch's objection, then, is that it runs the risk of overgeneralizing to many human activities and projects that are uncontroversially worthwhile. (Koch 2020b, 8)

I want to briefly consider another worry for Koch's approach. Even though one might not be required to have *antecedent knowledge* about the outcome, one might still be required or at least interested in having a *consequent knowledge* about the outcome. One could still argue that it is important to have *consequent knowledge* or knowledge that the implementation has succeeded. Here is an expansion of the above mentioned example from Deutsch. Imagine a case in which semantic change

[124] For more details see, Cappelen 2018, 72–73.

[125] Cf. also Riggs (2019) for why conceptual engineers should not worry about semantic externalism.

occurs as a result of us exercising our collective long-range control. As a result, *t* ends up meaning something else than we wanted it to mean, due to a mistake we have made (sufficiently) collectively. However, given that we were not aware of it, we kept believing that we have succeeded. If Koch's picture is right, then no matter how important, it would still be very hard (or sometimes even impossible) to have *consequent knowledge about an outcome*. Consequent knowledge about an outcome can play a significant role, both theoretical as well as practical. In particular, it can be important for the *re-evaluation* stage (see Section 2.3.2.5) of the *five stage recursive model* because having consequent knowledge about the outcome of previous stages is often a pre-condition to start re-evaluation as well as to determine the course of further re-engineering, or to determine whether re-engineering is needed. In some cases, this might also, thus, prevent us from deploying the recursive feature of my model. Furthermore, consequent knowledge about the outcome can be important for projects where precise semantic engineering is crucial, as opposed to some other projects where e.g. pragmatic engineering will suffice (see Chapter 6).

Here is my quick response to this worry. As mentioned earlier, for the purposes of building my framework, I want to stay neutral on the correct metasemantics. However, if externalism of this sort is true, then I am ready to bite the bullet and admit that this strategy, e.g. a very precise semantic engineering under externalism, even if possible, may face serious epistemic limitations. Furthermore, it is also worth stressing that, according to Deutsch,[126] the implementation feasibility worry concerns only *stipulative revision of semantic reference and meaning*. In my defense, however, this is only one of the strategies that my framework supports, i.e. my framework has other strategies at its disposal. Recall, my framework endorses pluralism both with respect to philosophical objects that can serve as FOP objects and pluralism about strategies one can employ to alter these objects. This means that this worry could affect my framework only partly in the case of externalism (if at all), namely only in cases in which one needs to employ the strategy of *stipulation* and take *semantic meanings* as FOP objects. Finally, according to my framework, to circumvent this unwanted result, at least sometimes, one could achieve (partially and/or temporarily) one's target result by applying some other strategy on a different object of engineering than semantic meaning. After all, semantic change is often not an end in itself, but is rather a tool to achieve some other change like practical, social, moral, etc.

The Feasibility Challenge for the Internalist about Semantic Meaning

An initial comparison between internalism and externalism with respect to conceptual change brought out by Burgess and Plunkett (2013a) suggests the following:

[126] It is, however, worth noticing that Deutsch (see 2020a, 6) does not deny the possibility of a semantic change.

> The textbook externalist thinks that our social and natural environments serve as heavy anchors, so to speak, for the interpretation of our individual thought and talk. The internalist, by contrast, grants us a greater degree of conceptual autonomy. One salient upshot of this disagreement is that effecting conceptual change looks comparatively easy from an internalist perspective. We can revise, eliminate, or replace our concepts without worrying about what the experts are up to, or what happens to be coming out of our taps. From the externalist's point of view, however, conceptual revolution takes a village, or a long trip to Twin Earth. (Burgess and Plunkett 2013a, 1096)

One might, thus, think that when it comes to semantic internalism it seems that an internalist should have better prospects of maintaining the *Naïve View of Implementation* (see Chapter 2, Section 2.3.2.4). However, Cappelen (2018) argues that the above mentioned Inscrutable and Lack of Control principles are not consistent only with externalism but also with internalism. Cappelen's (2018, 82) challenge to internalists relies on the fact that both externalism and internalism commit to a supervenience claim.[127] This leads to three issues that an internalist needs to deal with, which Pollock (2020) summarizes as follows:

> The first issue is that, even if meaning supervenes solely on states of the individual, it might supervene on states that we do not have epistemic access to. Secondly, even if it supervenes *only* on introspectively accessible states, these might be states over which an individual has no control. Lastly, even if these states themselves are both introspectively accessible and under the individual's control, it may be that meaning supervenes on these states in unstable or unpredictable ways such that we do not know (and thus cannot control) how changes to these states will affect the meaning of an expression. (Pollock 2020, 6)

However, against Cappelen's challenge to internalism, Pollock (2020) defends a strand of internalism she calls 'internalist conceptual role theories' by providing arguments for the following three claims that Cappelen's challenge invites internalists to meet: (i) there are inner states that are scrutable and within our control; (ii) meaning supervenes on these inner states; and (iii) the determination relation from supervenience base to meaning is scrutable and within our control (see Pollock 2020, 17).

Nevertheless, regardless of whether internalism is true and regardless of whether it can answer Cappelen's challenge, my general response is similar to the one I offered to the externalist worry (see above). In a nutshell, the negative result of a possible epistemic limitation would affect only part of the philosophical method behind UA, given that, according to my view, the philosophical method behind UA is broader and includes more than just semantic engineering. Moreover, in some cases, other subkinds of philosophical engineering may be able to compensate for this deficiency, especially when the main target is not engineering semantic properties.

---

[127] "… even if there's supervenience on what we want or intend or decide, the supervenience relation doesn't have to make it the case that semantic values are what we intend for them to be, what we want them to be, or what we agree on them to be (for all we know, it could be a total mess or get us to the opposite of what we want, intend, or decide)." (Cappelen 2018, 82)

The Implementation Value of Stipulative Revision of Speaker's Reference and Meaning

Recall, the second horn of Deutsch's sub-dilemma does not deny implementation feasibility for the speaker-meaning picture. Rather, it strongly *questions its value*. Deutsch claims that taking speaker-meanings as the *right* philosophical object behind UA trivializes the method behind UA:

> Perhaps Haslanger has changed what she and some of her readers speaker-mean by 'woman', and perhaps a conceptual engineer might convince some speakers to speaker-refer to something other than the semantic reference of 'free action' when using the term. But this seems like a rather trivial and easy thing to do. Surely it is not the sort of thing the exciting terminology of "conceptual engineering" was designed to describe. (Deutsch 2020a, 7)

However, Deutsch (2020a) neglects some important benefits of the speaker-meaning picture. For instance, Pinder (2019), who defends the *speaker-meaning* account, argues that his picture has two key benefits: it is often easy to achieve in practice and it offers a normative response to the objection that the method behind UA can only change the subject (see Pinder 2019, 1). Moreover, even though the speaker-meaning picture endorses a version of a *Naïve View of Implementation*, Pinder (2019) believes that there are some constraints that make it less trivial:

> It does not follow that we can speaker-mean anything we like by any term whatsoever: our expectations of the audience may be affected by the audience's linguistic sophistication, the complexity of the definition, the linguistic context, the political and social climate, and so on. But if the speaker is explicit about her intentions then, in an appropriate context, with an appropriate audience, speaker-meaning can deviate from semantic-meaning in new and substantial ways, and the audience will be able to grasp that speaker-meaning. (Pinder 2019, 13)

If so, the implementation of a change in speaker-meaning is not as trivial and easy to do as Deutsch claims it to be.[128] According to my framework, the speaker-meaning change, which takes speaker-meaning as one of the FOP objects, is a valuable subfield of the method behind UA given its suitability and sufficiency for certain kinds of FOPs. In particular, in order to achieve the desired change, in certain cases, it will be more suitable to operate on speaker-meaning (or some other non-semantic elements[129]) than on semantic meaning. Moreover, in some cases this kind of change will be sufficient to achieve the desired change.

I also want to consider another potential objection to the speaker-meaning account, or pragmatic accounts more generally—the *longevity* of the implementation change. Under the assumption that a permanent change is often more desirable, one may argue that changing the speaker-meaning of

---

[128] For more recent pros and cons of understanding the method behind UA in terms of the speaker-meaning, see Deutsch 2020b and Pinder 2020.

[129] In Chapter 6, I put forward another pragmatic account which, instead on speaker-meaning, operates on speech acts and propositional attitudes towards generic propositions.

MARRIAGE is less permanent and, thus, more inferior than changing the semantic meaning of MARRIAGE. This does not, however, present an issue for my model. Given that I am a pluralist about FOP objects, this means that my framework has the means to potentially meet the need by having other, more permanent, strategies at its disposal, e.g. in cases where change needs to be long-lived. Furthermore, having a strategy that will give short term results does not always need to be a bad thing. This will often depend on different goals one may have. For example, when one needs a short-term change in meaning, one might turn to changing speaker-meaning instead of semantic meaning, and turn to semantic meaning (or some other, more permanent) change when the more permanent change is required. Furthermore, speaker-meaning change is also a better option than semantic meaning change when the meaning change needs to be not only short-lived but also limited in scope, namely when the meaning change targets are limited to a particular audience, as opposed to the linguistic community as a whole.

The Hard Problem of Implementation

Consider a version of a *Naïve View of Implementation* that applies to philosophical objects such as concepts and word meanings. Pollock (2020) calls this kind of view the '*Autonomy View*': "on this approach, concepts and word meanings are entities that agents can deliberately manipulate in stable and predictable ways" (Pollock 2020, 2). Pollock (2020, 2) takes the Autonomy View to focus on "the question of whether concepts can be deliberately created or revised in predictable ways." She sets "the issue of successful implementation of revised concepts (in either an individual or a community) to one side." In this sense, the Autonomy View, according to Pollock, is "concerned specifically with the issue of conceptual revision or creation, rather than the further issue of implementation" (Pollock 2020, 2).

Pollock's distinction helps us refine the *implementation* problem from the *creation* problem. Consequently, it helps us realize that externalism and internalism are neither the main nor the only obstacles to implementation. The main challenge for the method goes beyond potential limitations imposed by a particular metasemantics. Moreover, one can generalize this distinction to any philosophical object behind UA. Given the differences between philosophical objects as well as possible strategies, it is safe to assume that the feasibility level of implementation in either an individual or a community could vary. Even when a particular strategy can be achieved in 'lab' conditions, one is still left with the task of maintenance and distribution of a particular strategy to its target audience. For instance, we may achieve a speaker-meaning change in 'lab' conditions, but distributing it to either an individual or certain (parts of) communities is, arguably, a more difficult project.

The *hard problem for implementation* of the method behind UA, thus, lies in implementation of a strategy operating on a particular philosophical object in either an individual or certain (parts of) communities. More generally, the worry about implementation feasibility can be re(de)fined into the

implementation feasibility of the following interconnected tasks:

(i) *Viability*: Is a particular strategy for the engineering of a particular philosophical object *possible*?

(ii) *Distribution:*[130] Can this change be successfully *transmitted* to the target audience, e.g. in either an individual or a community?

(iii) *Longevity*: Can this change, if needed, be successfully *maintained* in either an individual or a community?

Even though my framework does not entirely avoid the *hard problem of implementation*, I believe that my framework which takes into consideration the *pluralist approach to philosophical objects* offers a more optimistic picture about implementation feasibility of the method behind UA. Having more objects to operate on at one's disposal can also enhance the chance of its implementation viability, distribution and longevity. For instance, often, there will be different strategies available that one may employ in order to achieve a certain aim. Even though a certain general strategy (see Section 2.3.2.3) may be *more optimal* to apply to a certain philosophical object, one may rather choose the one that is *more implementable*. For example, in certain cases, it may be more needed and optimal to achieve *semantic meaning* change, given that a meaning change, once achieved, is more permanent than e.g. s*peaker-meaning* change. However, sometimes, given e.g. time and resources one has, all one can achieve on a large scale may be e.g. *speaker-meaning* change.

Finally, I want to briefly stress two compatibility features of my framework. First, my framework is largely compatible with a more general view on which *natural selection* and *rational reflection* utilized in philosophical engineering can go hand in hand. For example, upon some *rational reflection* during the strategic planning stage, certain strategies get proposed and dropped into the 'philosophical gene pool.'[131] Even though we cannot control or predict the process of natural selection, rational reflection can influence natural selection by facilitating the process of passing on strategies with better *fitness* (for more details see Chapter 5, Section 5.4.1.3, and Chapter 6, Section 6.2.4). Second, my framework is compatible with *outsourcing strategic planning and/or implementation stage* (either in part or entirely) of the

[130] When it comes to distribution, it is worth distinguishing between the *local* and *global scope* of implementation. In the former case, a proposed strategy can be applied to a local domain, i.e. within small or expert communities. For example, in the case of concepts, consider Scharp's (2013) suggestion to conceptually engineer TRUTH by splitting it to two concepts ASCENDING TRUTH and DESCENDING TRUTH for the purposes of logic and mathematics but keeping TRUTH in everyday communication. On the other hand, the proposed strategy can be applied globally, or to all possible domains of application for a philosophical object in question. For example, consider Haslanger's (2012) case of amelioration of WOMAN.

[131] Thanks to Patrick Greenough for suggesting this term.

method behind UA to non-philosophical methods and engineerings. And, *vice versa*, i.e. sometimes philosophical engineering will be in service of some other, non-philosophical engineering (see Section 2.3.3).

4.2 Addressing Worries about the Terminological Dispute Resolution

In Chapter 3, "Terminological Dispute Resolution of the Unity Assumption," I offered a novel resolution of the *terminological* dispute about the philosophical method behind the Unity Assumption (UA). First, I introduced a general definition of a *terminological dispute* and, as a case study, I suggested a terminological dispute about the philosophical method behind UA. Second, I evaluated two prominent terminological choices, i.e. 'conceptual engineering' (see Scharp 2013, 2020; Eklund 2015; Cappelen 2018; Burgess, Cappelen and Plunkett 2020, Chalmers 2020) and 'conceptual ethics' (see Burgess and Plunkett 2013a,b; Plunkett and Sundell 2013; Burgess 2014; Burgess and Plunkett 2020), against the two main conditions that I suggest one should take on board when making terminological choices for the philosophical method behind UA: (i) *the semantic correctness condition* drawing on Belleri (2018), and (ii) *the beneficial lexical effects condition* drawing on Cappelen (2018) and Chalmers (2011, 2020). Finally, I engaged in a *terminological dispute resolution* by arguing that we should introduce a new term, 'philosophical engineering.' For the philosophical method in question this terminological choice, I argued, fares better against its competition with respect to both of the above mentioned conditions.

In this section, I address five types of objections related to this chapter: Semantic Correctness or Descriptive Adequacy (Section 4.2.1); Aptness to Cause Misunderstanding, Miscommunication, or Harm (4.2.2); Too Broad and Too Narrow Terminological Choice (Section 4.2.3); What 'Philosophical' Means is Just as Unclear as What 'Conceptual' Means (Section 4.2.4); No Decisive Reasons for Introducing 'Philosophical Engineering' (Section 4.2.5).

4.2.1 Semantic Correctness or Descriptive Adequacy

The objector may suggest that Cappelen's advocacy of 'conceptual engineering' may not be, in Belleri's (2018) sense, *faulty* since one may argue that Cappelen's suggested use of 'conceptual engineering' does not deviate from how that term is used in the relevant language. To elaborate, the relevant language in question is 'Philosopher's English.' And it seems that, insofar as there *was* a prior standard usage of 'conceptual engineering' in Philosopher's English, Cappelen was adopting it. He follows, for example, the usage of Blackburn, Eklund and Scharp. Relevant here is that, on Cappelen's usage, 'conceptual engineering' is not a descriptive term. Rather, it seems to be defined by ostension, quite explicitly conforming extensionally to prior usage. Thus, in Belleri's sense, Cappelen's use of 'conceptual engineering' is semantically correct.

My response is that Cappelen's use does deviate from how that term was used in the relevant language and that Cappelen was not adopting the standard usage of the term in the relevant language. Blackburn, Eklund and Scharp use the term 'conceptual' in 'conceptual engineering' *descriptively*. According to them, FOPs are about engineering *concepts*. Prior to Cappelen, in the relevant language, the use of 'conceptual' in 'conceptual engineering' has, thus, been descriptively adequate which translated to the semantic meaning and the content behind the term being about concepts. However, both of Cappelen's uses of the term 'conceptual engineering,' i.e. the one for a general methodological field that Chapter 3 focuses on as well as the one concerning its subfield that Cappelen spells out under the name Austerity Framework, are neither *descriptively adequate* nor *semantically correct* but are, at best, only partially so. In comparison to Blackburn, Eklund and Scharp's original use of the term, when speaking about the general methodological field Cappelen broadens the standard meaning behind the term 'conceptual engineering' to include other FOP objects than concepts only, whereas in the case of his own Austerity Framework he narrows it down to intensions.

The objector might further argue that for my purposes it does not really matter either way given that my argument seems to turn on the claim that Cappelen's use of 'conceptual engineering' is *descriptively inadequate*—a claim that seems to hold regardless of whether the comments above are right.

The objector's point, as I take it, boils down to whether it makes sense to count the semantic correctness condition at all when making our terminological choices. As long we have good reasons to value things such as e.g. truth, understanding and communication in the relevant language,[132] I am inclined to say yes, at least in majority of the cases. For example, a semantically correct use of an expression conforms with the expression's meaning and content and it, thus, contributes to the truth of the proposition in which it occurs. Furthermore, when an expression conforms with the expression's meaning and content, this puts the audience in a better position to understand the meaning and the contents of the proposition in question. Finally, this also contributes to a successful communication of that very proposition.

However, the objector may proceed and say that, given my *first* response, i.e. that whether Cappelen's use of 'conceptual engineering' meets the semantic correctness condition depends on whether Blackburn, Eklund and Scharp used the term descriptively, call that (*), and, given my second response, i.e. my claim that meeting the semantic correctness condition is important for communication and understanding, the initial thought of the objector may be restated as follows. The arguments against Cappelen's use of 'conceptual engineering' that I offer are completely independent of how Blackburn, Eklund and Scharp used the term. Given those arguments, it seems that whether Cappelen's use of

[132] In Chapter 6, however, I argue that in the case of certain expressions such as generics (given the nature of their semantics and truth-conditions), our moral, political and social values can trump the epistemic and semantic ones.

'conceptual engineering' will lead to misunderstandings or miscommunications hasn't got anything to do with whether Blackburn, Eklund or Scharp used the term descriptively. So, in this particular case, given (*), it does not seem that communicative success and understanding have anything to do with the semantic correctness condition. Rather, it has to do with *descriptive inadequacy*. The main objection is, thus, not about my arguments here—but that I would make my argument more convincing by arguing directly that Cappelen's use is problematic because it is descriptively inadequate, rather than going via the semantic correctness condition. In that sense, the discussion of faulty vs. faultless verbal disputes might be a bit of a red herring.

My response is that I disagree that, in this particular case, given (*), it doesn't seem that communicative success and understanding have anything to do with the semantic correctness condition. For Belleri (2018), semantic correctness is about conforming to what, on the basis of usage (of a certain language L), are considered the expression's meaning and content. To establish that Cappelen's use is semantically incorrect in that sense, one needs to establish that he is not conforming to what, on the basis of usage in Philosopher's English, are considered the meaning and content of 'conceptual engineering,' as I have suggested above. Even though descriptive adequacy does not need to match semantic correctness condition, in this particular case it does given that Blackburn, Eklund and Scharp's use is descriptively adequate. In other words, in the relevant language, i.e. 'Philosopher's Language,' in order for Cappelen's use of the term 'conceptual engineering' to be descriptively adequate, it has to also be a semantically correct use, given that the standard meaning behind Blackburn, Eklund or Scharp's terminology were descriptively adequate too.

Finally, it is important to notice that it is also likely that Cappelen is not deferring to Blackburn, Eklund or Scharp's terminology ostensively, given what he takes the term 'conceptual engineering' to stand for, both when he uses it for the philosophical method behind UA as well as when he uses it for his Austerity Framework. Instead, Cappelen can be seen as only *borrowing* the term from them because of the lexical effects of the term, rather than because he thinks that Blackburn, Eklund or Scharp's use of the term is semantically correct or descriptively adequate. After all, Cappelen does not seem to care much about the descriptive adequacy of the term and he points this out in several places. Moreover, not only does he not care if Blackburn, Eklund or Scharp are using it descriptively, he himself does not care when he is not using it descriptively. Thus, even if, according to Cappelen, the term 'conceptual engineering' is supposed to function as a proper name rather than a description, and he does not defer to Blackburn, Eklund and Scharp's use of the term but he rather only borrows the term, one may still argue that this terminological choice can have misguiding or bad semantic and lexical effects in the similar way in which Cappelen's example of the baby named Hitler has (see Chapter 3, Section 3.4).

4.2.2 Aptness to Cause Misunderstanding, Miscommunication, or Harm

One may argue that the mere possibility for misunderstanding, miscommunication, or harm does not seem to be enough to outright reject the use of words that do not conform to standard linguistic meaning. Should not one have to show that a *particular* use of a term is *apt* to cause misunderstanding, miscommunication, or harm?

The main question, as I understand it, boils down to an asymmetry or difference between the *mere possibility of something* vs. *something being apt to cause.* To this, one could perhaps also add something stronger like vs. *something will cause* vs. *something causes,* etc. It is true that there is a matter of degree there that may carry different weight when providing justification for our terminological choices. However, if one also considers other factors such as the *costs* or *consequences* of misunderstanding, miscommunication, or harm for a particular case, it might be more justifiable to reject the use of certain terms (that do not conform to standard linguistic meaning or do (not) bring about certain lexical effects) even if there is less of a certainty, especially when the stakes are high. For instance, in certain contexts, when the costs are high, one might want to try to prevent certain misunderstanding, miscommunication, or harm even if one deems that there is only a possibility that it is going to cause misunderstanding, miscommunication, or harm. In the present context, which is highly theoretical, I suggested that 'conceptual engineering' and 'conceptual ethics' are at least *apt to cause* misunderstanding, miscommunication, or harm, and that, despite not being ideal, 'philosophical engineering' fares better in that respect, and should, thus, be considered as a terminological choice.

Relatedly, the objector could say that even despite being apt to cause misunderstanding, miscommunication, or harm, one should also have to show that there must be an absence of *feasible means of avoiding* miscommunication, misunderstanding, or harm with the use of a term (i.e., by carefully stating what one means, such as Cappelen does when he lays out the Austerity Framework). So, while there might be misunderstanding or miscommunication (perhaps not harm), such can be corrected with further discourse or meaning change.

As I briefly point out in Chapter 3, I believe that the semantic correctness condition may be long term compatible with different kinds of philosophical engineering. For instance, there may be cases in which in order to do e.g. semantic engineering one needs to go against *current* semantic correctness conditions in order to change the meaning of a certain concept. Once the meaning is changed things go back to normal and the semantic correctness gets reestablished. In such cases, one may argue that reasons that are sufficient for meaning change may be sufficient to temporarily trump the semantic correctness condition. To put this in the present context, one may argue that what Cappelen is doing or what I am doing is trying to change the meaning behind the terms 'conceptual engineering' and 'philosophical engineering' in the relevant language. Even if so, I still believe that the term

'philosophical engineering' may be more apt (to be ameliorated) for the purpose in question than its current competitors. On the other hand, I would argue that Cappelen does not have better reasons to change the meaning behind the term 'conceptual engineering' to cover FOP objects other than concepts. Even if the semantic change occurs and 'conceptual' in 'conceptual engineering' stands for FOP objects other than concepts, as his substantive dispute resolution of the philosophical method behind UA suggests, the term 'conceptual' will still be more inadequate than the term 'philosophical' with respect to potential lexical effects.

Relatedly, the objector could worry that it seems that we cannot develop general principles for the use of a term since much of what makes misunderstanding, miscommunication, and harm likely is context-sensitive.

I agree that there are many obstacles for giving a general theory, including the one mentioned above, namely that misunderstanding, miscommunication, or harm likely are context-sensitive. However, as I briefly point out in Chapter 3, Belleri (2018) believes that her distinction could also be applied to non-literal uses. She spells it out in terms of *pragmatic correctness*. As she points out:

> Considering non-literal uses as well would probably require pragmatic correctness—definable as a use's successful communication of a content that is reasonable given the context of utterance. Even with pragmatic correctness, we could still distinguish between merely verbal disputes that are faultless (because both uses are pragmatically correct) and merely verbal disputes that are faulty (because at least one is pragmatically incorrect). (Belleri 2018, 696)

In certain contexts, where non-literal uses are important for avoiding misunderstanding, miscommunication, or harm, paying attention to pragmatic correctness could help. In connection to that, I also want to point out that the particular cases we are looking at are situated in a highly theoretical context where pragmatics and context-sensitivity play at least somewhat lower role than in some other, highly context-sensitive cases such as e.g. poetry or music.

However, looking only at semantic or pragmatic correctness when making terminological choices will not always suffice to avoid misunderstanding, miscommunication, or harm given that there are other non-semantic and non-pragmatic factors that may contribute to misunderstanding, miscommunication, or harm, such as lexical effects. Offering general principles for lexical effects of our terminological choice requires further theoretical and empirical work.

4.2.3 Too Broad and Too Narrow Terminological Choice

On the one hand one might wonder if my terminological choice 'philosophical engineering' is not *too broad* in the sense that the term 'philosophical' does not pick out the specific things that are being engineered.

Given my substantive dispute resolution about the method behind UA (see Chapter 2), I take 'philosophical engineering' to stand for the substantial debate which is more general than e.g. terminological engineering, pragmatic engineering, semantic engineering, intension engineering, etc., all of which may be seen as its possible subfields. However, even if too broad, I believe that the term 'philosophical engineering' could change its meaning making it, thus, more or less narrow (for a response to a similar worry related to the substantive dispute resolution, see also Section 4.1.1).

On the other hand, one might wonder if my terminological choice 'philosophical engineering' is *too narrow* given that the term 'philosophical' *sounds* prohibitively narrow in its implication that the philosophical method behind UA is uniquely a philosophical endeavor.

Even if it is true that the term 'philosophical engineering' makes the method of engineering in philosophy sound as a distinctively philosophical method, I take this to be a positive thing. I addressed the worry about the method of philosophical engineering not being a uniquely a philosophical endeavor in Section 4.1.2. Furthermore, in Chapter 2, Section 2.3.3, I advocated for *engineering resemblance*, *division of engineering labour* and *engineering holism*.

Moreover, I believe that there is nothing prohibitively narrow in having a term for the general philosophical method, especially since the term and the method, I argued, can be seen on a par with other subkinds of engineerings such as e.g. chemical engineering, mechanical engineering, or software engineering. Interestingly enough, however, one does not usually apply the same kind of worry to the latter kinds engineerings despite these kinds of engineerings also being done not only by e.g. chemical engineers, mechanical engineers, or software engineers, respectively.

Finally, one might argue that the term 'philosophical engineering' could make the method less attractive (and maybe heavily loaded with bad connotations as well) to people outside of philosophy.

I believe that, on the contrary, with help of more discussion and shared knowledge about the general philosophical method behind UA as being about *engineering* and being *philosophical*, the term 'philosophical engineering' and the meaning behind the term could contribute to changing non-philosophers' negative attitudes towards philosophy.[133] Compare, for instance, the difference in lexical effects of the term 'philosophical analysis' as opposed to 'philosophizing' when answering the question what are philosophers *doing*.

4.2.4 What 'Philosophical' Means is Just as Unclear as What 'Conceptual' Means

One might have some reservations about my terminological choice 'philosophical engineering' in the sense that what 'philosophical' means seems to be unclear just as what 'conceptual' in 'conceptual

[133] However, measuring as well as potentially changing the lexical effects of a particular piece of terminology are to a large extent empirical matters.

engineering' or 'conceptual ethics' means. One might argue that what we call 'philosophy' refers to a bunch of heterogeneous interests, methods and problems that cannot really be captured with a single definition. The question there is whether I am assuming that there is some sort of clarity about the meaning of the term 'philosophical'?

In response, I am ready to accept this potential unclarity as a reservation about my terminological choice of 'philosophical' in 'philosophical engineering.' However, it is worth noticing that I did not argue that my term is optimal or connected to an optimal semantics or an optimal concept with optimal lexical effects. Rather, I aimed to offer a term that, I believe, has a potential to fare better than its competition for the philosophical method behind UA. For instance, to compare, given that current debates on the topic often use the same term 'conceptual engineering' to discuss different substantive questions in the vicinity, the term 'conceptual engineering' is either homonymous or polysemous, i.e. it used for the general method behind UA as well for different potential sub-methods, including engineering concepts but also other objects. On the other hand, the term 'philosophical engineering' that I suggest is supposed to be used only for the general philosophical method behind UA whereas other terms should be used for its potential sub-methods. It is, thus, worth noting that I believe that the term 'philosophical engineering' is better than 'conceptual engineering' or 'conceptual ethics' only in cases in which we want to refer to the method of engineering in philosophy *in general* rather than to its particular sub-methods. In other words, I am not opposed to using terms like e.g. 'conceptual engineering,' 'semantic engineering,' or 'terminological engineering,' etc., when we want to refer to engineering concepts, semantic properties, or terms, etc. Moreover, even if what we currently call 'philosophical' refers to a bunch of heterogeneous interests, methods and problems, this does not mean that it could not or should not change its meaning in order to serve our theoretical purposes better.[134] Even though the terminology I offered might be, at the moment, unclear in a certain sense, I believe that the term 'philosophical engineering' has the potential to be (epistemically) clarified or semantically changed if needed, which I believe is compatible with semantic correctness conditions (see Section 4.2.1).

I would also like to add that apart from the unclarity about the meaning of the term, sometimes the terminological unclarity can be due to different effects connected to the term such as e.g. lexical, semantic, practical, or cognitive ones, rather than (just) the meaning behind the term itself. For instance, from a semantic point of view, in order to refer to e.g. the concept MARRIAGE it is possible to imagine that, in certain contexts, one could be able to semantically correctly use (or, perhaps, stipulate) different terms such as 'marriage' or 'darriage' or 'mrg,' etc. However, the lexical effects of the terms e.g. 'marriage' vs. 'darriage' vs. 'mrg' may be different, a reason why one currently prefers 'marriage' over

[134] See engineering of conceptual engineering with e.g. Cappelen (2018); Chalmers (2020); Isaac (2020).

the others. Similarly, even if, in certain contexts, or due to a semantic change, one may manage to use the term 'conceptual engineering' semantically correctly for the general philosophical method behind UA, the effects of using the term 'conceptual engineering' instead of 'philosophical engineering' can be different. In Chapter 3, I suggested that, even if so, the term 'philosophical engineering' may still have better effects than the above mentioned alternatives. Relatedly, given that the substantive dispute I focus on is the one about the general philosophical method which is, in its generality and structure, more similar to philosophical analysis than to its subfields, I believe that the structural and cognitive parallels with philosophical analysis can make the effects of the term 'philosophical engineering' more beneficial than the effects of the term 'conceptual engineering.'

The objector may, however, argue that there is a disanalogy to the case of 'philosophical analysis.' One would think that 'philosophical' in 'philosophical analysis' neither means that the analysis is done by a philosopher, nor that that which is being analyzed is a philosophical object. Rather, 'philosophical' modifies 'analysis'—the analysis is done in a particular way, namely a philosophical way. What that means exactly is a big question. Presumably, many thought it involves stating definitions, considering thought experiments and taking into account the intuitive case verdicts they generate, etc. One may, thus, think that there is a clearer sense in which an analysis is done philosophically than a sense in which certain objects are philosophical.

I disagree that there is a clearer sense in which an analysis is done philosophically than a sense in which certain objects are philosophical. I believe that the same worry applies to "being done in a philosophical way," given that other disciplines and sciences utilize the same strategies that have been suggested above, i.e. they state definitions, consider thought experiments and take into account the intuitive case verdicts they generate, etc. On the other hand, I argued that engineering philosophical objects makes the philosophical method behind UA philosophical. However, as I mentioned above, I am inclined to think that there is a sense in which none of the objects is strictly a philosophical object other than getting attention from philosophers (see Section 4.1.1) and that neither analysis nor engineering have to be done entirely by philosophers (see Section 4.1.2). Moreover, as I also mentioned, what a philosophical object is can change and has changed throughout history, especially with advances in other disciplines and sciences and their interests in objects that have once been philosophical objects. In the end, even if one disagrees whether it is *philosophical objects* or *philosophical ways of doing things* that make the method behind UA *philosophical*, one may still be justified to call the method 'philosophical' given that those working in philosophical methodology are trying to demarcate a method that is philosophical.

4.2.5 No Decisive Reasons for Introducing 'Philosophical Engineering'

The objector could argue that all of the different terminological choices that e.g. Cappelen, Plunkett and Burgess thought about are ones that are suboptimal in various ways, given the relevant normative/evaluative criteria that matter in the relevant contexts. One may think that is certainly true of the term 'conceptual ethics,' as well as 'conceptual engineering.' However, one may doubt that these terms are, all-things-considered, more sub-optimal than the relevant alternatives—including my terminological choice 'philosophical engineering.' The tests I am bringing in are providing some reasons for favoring certain views over others. But there is no great argument that those reasons should necessarily be decisive, given other considerations at play. Given that, one may ask how strong do I think these reasons are? Are they meant to be *decisive*? Or just strong reasons that motivate looking for a new term that might do better? In other words, what are the consequences that I think we should draw from my argument? Do I merely mean that 'philosophical engineering' would be a better term than e.g. 'conceptual engineering'? (As in: the world would have been a better place if Cappelen et al. had chosen that term). Or do I think we should swap terminology now and start using 'philosophical engineering'?

I believe that the reasons I offer are not decisive but rather, as suggested above, motivational enough to look for a better term. Furthermore, the proposal of introducing the term 'philosophical engineering' follows from the results of the first two chapters of Part I of this dissertation. In a nutshell, Part I of the dissertation has three main aims: to identify the phenomenon and different disputes (i.e. merely verbal, substantive and terminological dispute) about the phenomenon of the general philosophical method behind UA (Chapter 1), to offer a substantive dispute resolution (Chapter 2) and terminological dispute resolution (Chapter 3) for the general philosophical method behind UA. In Chapter 3, on the one hand, my aim was to show that the current terminology for the general method behind UA is not ideal with respect to the two main conditions I believe one should consider when making terminological choices, especially in highly theoretical cases: i.e. *the semantic correctness condition* and *the beneficial lexical effects condition*. On the other hand, given the above two conditions, I suggested 'philosophical engineering' as a better competitor than the current alternatives. Within this context, my claim was that the term 'philosophical engineering' *would be* a better term than e.g. 'conceptual engineering' for the general philosophical method behind UA. In this sense, my proposal can be seen as a call for relabelling. Furthermore, I believe that this relabelling is an important one, especially since, I believe, it has a greater potential than current alternatives to, on the one hand, help us avoid the confusion and hindrance of epistemic goals and, on the other hand, help us see the phenomenon behind FOP in a deeper and more unified way.

However, the objector could argue that if I am really advocating the use of a different terminology now, then I would have to take into account many other factors as well. There is a huge debate with

100+ publications that is arranged around the terminology of 'conceptual engineering' and, to a lesser extent, around 'conceptual ethics' already. This, one may argue, gives us very strong reasons to keep the terminology. After all, the label 'conceptual engineering' is immensely important for coordinating our activities, flagging relevance to each other, attracting attention, etc. Even if I am right and it would in some sense have been better to go by 'philosophical engineering,' it seems to be a completely different question whether this entails that we should swap terminology now. Even if the term 'philosophical engineering' is more optimal than the alternatives with respect to the above mentioned conditions, there is also a question of whether it is the optimal terminological choice that will have a sufficient uptake in the relevant community.

My brief response is that I do not believe that it would be necessarily *so* difficult to implement into the relevant community the introduction of the term 'philosophical engineering' for the general philosophical method behind UA. After all, what I am proposing is not at all that new in philosophy nor in other parts of human inquiry. New theoretical terminology (in many cases alongside a new concept) is often being introduced to e.g. distinguish, refine, broaden or narrow down either the same phenomenon or the phenomena in the vicinity.[135] The uptake of the term is often limited to the relevant community even though it can in some cases also be popularized and spread among wider audiences. I believe that as those working in the field of the philosophical method behind UA recognize more and more the possibility that the scope of the methodological field in question can include other objects than just concepts and conceive different versions of engineering in philosophy not only as competitors but also, in some cases, as co-fields/co-workers (cf. Chapter 2), the chances of accepting into the relevant community the term 'philosophical engineering' for the general philosophical method would get even higher.[136] However, given that measuring the implementation success is an empirical issue, I do not take the term 'philosophical engining' to be the knock-down, obvious winner here, but I believe that long-term the term 'philosophical engineering' could lead to at least equally good uptake in the literature, especially when brought into connection with philosophical analysis and its structural analogy.

To conclude, terms are often crucial in demarcating and communicating certain phenomenon we deem important. The importance of having a term that demarcates and communicates *better* our

---

[135] On the other hand, the old terminology sometimes gets retained. As Chalmers puts it: "This was roughly our attitude in 'The Extended Mind'. Andy and I could have introduced a new term, 'e-believe', to cover all these extended cases, and we could have only made claims about how unified e-belief is with the ordinary cases of believing and how e-belief plays the most important role. We could have done that, but what fun would that have been? We thought e-belief plays the most important roles associated with 'belief', and that as a result 'belief' picks out e-belief. Saying that belief is e-belief helps to communicate those claims. Furthermore, in practice the word 'belief' plays a central role in both social and scientific explanation. So attaching the word 'belief' to e-belief helps to create a mindset where e-belief can actually play those roles" (Chalmers 2020, 8).

[136] Chalmers (2020) makes a similar point with respect to the term 'engineering': "Of course there is a legitimate narrower subject matter of conceptual re-engineering with some distinctive concerns. At the same time, words have power, and 'conceptual engineering' is a powerful expression. If people had just talked of 'concept repair' or 'concept re-engineering' from the start, I don't think the interest would have been so great. The power comes in part because of the suggestion of something that goes beyond mere fixing. I think we should use the term in a way that vindicates that power" (Chalmers 2020, 8-9).

theorizing about the phenomenon of the general philosophical method behind UA partly falls out from the importance of theorizing about the general philosophical method behind UA (see Section 4.1.1). Even though there is a degree of arbitrariness about which term will stand for which phenomenon, changing the terminology does usually not come easy, especially after a certain term has already been introduced and used for a certain phenomenon. There are different factors, including practical and cognitive, at play, which can make changing terminology difficult. For these reasons, it seems that one needs to consider extra justification for changing the terminology. In Chapter 3, I considered some *epistemic* (e.g. avoiding misunderstanding, miscommunication) and *practical* reasons (such as beneficial semantic and lexical effects) for choosing a particular term for the general philosophical method behind UA, which I formulated in terms of two conditions: *the semantic correctness condition* and *the beneficial lexical effects condition*. I also argued that epistemic reasons should often trump practical ones, especially in highly theoretical contexts. I evaluated current terminology for the general philosophical method behind UA as *inadequate* given these two conditions. I took the inadequacy of the current terminology to motivate the project of looking for a better term. The second part of Chapter 3 is thus an attempt to provide us a better term. The term I suggested to introduce was 'philosophical engineering.' However, I have not offered decisive reasons for introducing 'philosophical engineering' nor have I argued that the term I offered is ideal. I have simply argued that it fares better than the current alternatives in the above sense. Moreover, I have offered some predictions about its uptake in the relevant community, with awarenesses that the uptake of the term 'philosophical engineering' will to a large extent be a matter of an empirical question. As with most of the similar cases, I suspect that the uptake of the term will also largely depend on the further visibility of my proposal and the influence of those willing to adopt it.

Part II

Pernicious, Beneficial and Neutralizing Generic Judgments: A Case Study of Philosophical Engineering

Easy to accept, hard to refute is not
an epistemically or politically promising combination.

(Langton, Haslanger and Anderson 2012, 764)

CHAPTER 5

A DARWINIAN DILEMMA FOR REALIST THEORIES OF GENERICITY

Abstract

In this chapter, I bring to the foreground *judgments about generic propositions*. In her paper called "A Darwinian Dilemma for Realist Theories of Value" Sharon Street (2006) argues against *realist theories of value*. In this chapter, I consider the prospects of applying an analogous Darwinian dilemma for *realist theories of genericity* by offering a debunking argument against genericity drawing on a direct analogy with Street's (2006) paper. In particular, I focus on generic statements such as "A tiger is striped" or "Muslims are terrorists" and look at the relation between *generic judgments* and *independent generic truths* expressed by generic statements posited by the realist about genericity. This chapter makes a broader methodological point about the connection between realist theories about genericity and speakers' judgments about generic propositions, based on the assumption that generic judgments have been indirectly yet significantly influenced by *evolutionary forces*.

5.1 Introduction

Generic expressions or generics (e.g. "Tigers have stripes" or "Muslims are terrorists") are propositions that express generalizations but do not specify how many members of the kind have the property being expressed.[137] For instance, it seems that some generics such as "Dogs are mammals," in order to be true, require that all members of a natural kind dog satisfy the property of being mammals. On the other hand, generics such as "Mosquitoes carry the West Nile virus" for its truth seem to be rather permissive since very few actual mosquitoes (less than 1%) carry the virus. There are also generics such as "Tigers are striped" or "A lion has a mane" that seem to be semi-permissive since there exist some tigers that are stripeless and only adult lions have mane. Generics can, thus, express weak or strong generalizations, given that speakers accept certain generics as true even when not all of them seem to express strong generalizations.

Because of their different forms and peculiar features, generics have turned out to be rather difficult to analyze.[138] This has motivated some theorists, especially those who are inclined to provide a unified semantics for generics, to posit the existence of a unifying metaphysical phenomenon they believe that generics instantiate. They call this metaphysical property 'genericity.' According to genericists, the truths that generic statements express, thus, become independent generic truths or the truths about *genericity*.

Offering a theory of *genericity* has become an important element for those who aim at offering a unified semantics for generics. However, it should be underscored from the start that the aim of this chapter is neither to criticize any particular theory of genericity nor to offer a theory of the truth-conditions for generic propositions. Instead, this chapter makes a broader methodological point about the connection between *realist theories about genericity* and *speakers' judgments about generic statements*.

In particular, in this chapter, I look at the relation between generic judgments and independent generic truths expressed by generic statements posited by the realist about genericity. In a nutshell, based on the assumption that generic judgments have been indirectly yet significantly influenced by evolutionary forces, I offer a debunking argument against genericity by drawing on a direct analogy with Street (2006).

Here is how I proceed. In Section 5.2, I offer a focused overview of the state of the art of generics. In Section 5.3, I introduce a distinction between generic judgments and generic truths, and contrast them with judgments and truths about overtly quantified statements. In Section 5.4, I work towards

---

[137] For example, if asked "how many ravens are black?" one could reply "all [or some, or most, etc.] ravens are black," but one cannot felicitously reply with the generic "ravens are black" (Carlson 1977).

[138] "Difficulties include: (i) what is the logical form of generic sentences—i. e., are they kind-predications or quantified sentences? (ii) Do generics express or merely convey generalisations? (iii) What generalisation or class of generalisations do generics communicate?" (Sterken 2015, 1)

establishing the analogy between Street's dilemma for realist theories about value and my dilemma for realist theories about genericity. In Section 5.5, I develop a Darwinian dilemma for realist theories about genericity based on a direct analogy with Street's dilemma. In Section 5.6, I conclude with some further implications of a Darwinian dilemma for realist theories about genericity.

5.2 Generics

5.2.1 Syntactic Form

In natural languages, there is no special syntactic form reserved for generic sentences. Instead, generics take different syntactic forms, e.g. *bare plurals* such as "Tigers are striped," *indefinite singulars* such as "A tiger is striped," or *definite singulars* such as "The tiger is striped." However, it is important to notice that, apart from a generic reading, each of these syntactical forms can, on certain occasions, also be used to express existential claims.[139]

Furthermore, we can distinguish between two broad categories of generics: (i) *characterizing generics* or *I-generics,* and (ii) *direct kind predications* or *D-generics*.[140] The I-generics such as "Tigers are striped" seem to express generalizations about individual members of the kind (where particular members of a kind posses the predicated property), whereas the D-generics such as "The dodo is extinct" are taken to be singular statements that are directly about kinds (where the kind itself possess the predicated property).[141]

A large amount of literature (including this chapter) has mainly been devoted to *characterizing bare plural generics* such as "Tigers are striped" as they proved to be the most controversial, as opposed to *direct kind generics* such as "The dodo is extinct" which do not seem to be problematic when it comes to their truth-conditions since in order for a direct kind predication generic to be true it takes that all the members of the kind have to have the property being predicated to them (since the prediction amounts to the kind directly).[142]

On the other hand, *definite singulars*, e.g. "The coke bottle has a narrow neck" and *indefinite*

---

[139] Two tests are usually used for these purposes. The first one relies on the assumption that "the existential interpretation is upward entailing, meaning that the statement will always remain true if we replace the subject term with a more inclusive term. (…) If "tigers are on the lawn" is true, then "animals are on the lawn" must be true. However, "tigers are striped" is true, yet "animals are striped" is false. (…) " The second test concerns whether we can insert an adverb of quantification with minimal change of meaning (Krifka et al. 1995). For example, inserting "usually" in the sentences in (1a) (e.g., "tigers are usually striped") produces only a small change in meaning, while inserting "usually" in (1b) dramatically alters the meaning of the sentence (e.g., "tigers are usually on the front lawn"). (For generics such as "mosquitoes carry malaria", the adverb "sometimes" is perhaps better used than "usually" to mark off the generic reading)" (see Section 1.1, Leslie and Lerner 2016).

[140] In *I-generics* "I" stands for "indefinite" whereas in *D-generics* "D" stands for "definite" (see Krifka 1987).

[141] This widely accepted distinction follows Carlson's (1977) notion that bare plurals refer directly to kinds.

[142] E.g. in the case of a direct kind predication generic such as "The dodo is extinct" in order for the (species) of dodo to be extinct every member of the kind has to be extinct too. It has, thus, been widely agreed that this kind of generics do not allow for exceptions, namely all the members of the kind (e.g. 'dodo') in question have to have the predicated property (e.g. 'being extinct').

singulars, e.g. "A madrigal is polyphonic" (see Lawler 1973) have received very little attention, the former even less than the later. Perhaps the least attention has been given to *bare singulars*, e.g. "Gold is a metal" (see Ojeda 1991, Carlson 2011).

5.2.2 Logical Structure

Generics, as opposed to universally and existentially quantified sentences, do not contain any pronounced equivalent of an operator. Moreover, it is often considered that generics do not behave as typical quantified statements and cannot be simply paraphrased with determiners such as "all" or "some." Furthermore, they are also taken to differ from the statements containing Lewis's (1975) adverbs of quantification such as "usually" or "generally."

However, the dominant view is the one that posits a *covert* operator *Gen* (e.g., Heim 1982; Lawler 1972; Nickel 2016; Schubert and Pelletier 1989; Sterken 2016). For example, for I-generics in particular, some (e.g. Carlson 1989;[143] Schubert and Pelletier 1987) have introduced a two-place covert operator *Gen*, based on Lewis's (1975) notion of a tripartite structure of adverbs of quantification, consisting of a quantifier *Q*, a restrictor *R* (picking out the domain of *Q*), and the scope *S*, also called the matrix (picking out the properties attributed to *Q* elements of *R*). For instance, the logical structure of a generic sentence of the form "Fs are Gs" is true just in case Gen(x) [F(x)] [G(x)], e.g. "Blacks are violent" is true just in case Gen(x) [Blacks(x)] [violent(x)].[144]

Recently, alternative views (e.g. Leslie 2007, 2008, 2012; Carlson 2011; Liebesman 2011; Sterken 2015) that reject the quantificational nature of *Gen* have also been offered. For example, Leslie (2007, 379) argues that "despite appearances, generics are in no sense quantificational. The generic operator Gen is a variable-binding operator that is used to express generalisations, and yet is not a quantifier" (Leslie 2007, 379). One reason for that, she believes, lies in that our intuitions about how many members of the kind have to satisfy the predicated property in order for a statement to be true or false are reversed. Consider examples such as "Ducks lay eggs" taken to be true vs. "Ducks are female" taken to be false, and "Mosquitos carry the West Nile virus" taken to be true vs. "Books are paperbacks" taken to be false. Moreover, Leslie (2008, 2012) argues that the word "all" we find in universal statements is needed as an explicit instruction because our cognitive system has to deviate from its default way of operating (i.e. expressing generalizations via generics) in order to express universal generalizations. This is based on

[143] Carlson (1977) first introduced a monodic predicate operator "G."

[144] Generalizations mainly concern kinds. However, in some cases the term generics is extended to stand for statements such as "Mary smokes after work" that are sometimes referred to as "habitual" instead of "generic" (see Leslie and Lerner 2016). Habitual statements are usually also analyzed with *Gen* (see Lawler 1972; Schubert and Pelletier 1989; Krifka et al. 1995).

her hypothesis (Leslie 2007, 2008) that generics are cognitively fundamental generalizations[145] as opposed to the quantified generalizations that are more taxing and sophisticated and, thus, need an explicit instruction such as "all" or "some." It is worth noticing that this account clashes with approaches that analyze *Gen* in terms of its reduction to theoretically more tractable quantifiers, such as, for example, some of the possible worlds and normalcy-based approaches that analyze *Gen* in terms of universal quantifiers.

Another notable exception is Liebesman (2011) who argues against the existence and need for any operator at all. In particular, Liebesman (2011, 411) claims that: "Surveying the literature on generics reveals a large supply of complicated and interesting examples, and to go with the examples, a large supply of complicated and interesting semantic accounts of Gen." However, he claims that there are extant counterexamples to every analysis. He believes that the intractability of Gen is unsurprising. On his view: "Gen has proven intractable for a very simple reason: it doesn't exist. Generics do not have the tripartite structure that theorist take them to have." (Liebesman 2011, 411). Instead, he takes them to be kind predications.

5.2.3 Semantic Approaches[146]

Despite numerous theories, it has proven to be extremely hard to give a formal semantic analyses of generics. To enlist some: *possible worlds approach* (see Pelletier and Asher 1997), *relevant quantification approach* (see Schubert and Pelletier 1987), *indexical approach* (see Sterken 2015), *stereotypes and prototypes approach* (see Geurts 1985; Declerk 1986), *probabilistic approach* (see Cohen 1996). In addition, some of these approaches appeal to *non-monotonic logic* (see McCarthy 1986; Reiter 1987), *restriction of the domain of discourse* (see Schubert and Pelletier 1987; Pelletier and Asher 1997), *constraints on situations* (see Cavedon and Glasbey 1994), *indexicality of a generic operator* Gen (see Sterken 2015), *comparison with other kinds* (see Cohen 1996, 1999, 2004).

Over the past half a century, many semanticists have been rather unsuccessfully working on what may be called an 'alchemist project' of finding the correct semantic account for the unpronounced operator *Gen*. However, the current state of the art of generics teaches us that, despite the existence of many alternative accounts, it has been very difficult to give a plausible and unified semantic account of generics. Instead, some theories can explain better certain subclasses of generics but have difficulties

---

[145] "It is not clear on the face of it what one should make of the fact that these truth-conditionally puzzling statements are apparently so easily processed by young children— considerably more easily, it would seem, than their quantified counterparts. In Leslie (2007, 2008) I proposed these results could be explained (or rather at that time, I predicted these results, with the exception of the earlier data in Hollander et al., 2002, which was of course an inspiration for the hypothesis) if we suppose that generic sentences are language's way of letting us give voice to cognitively fundamental generalizations. We observe that infants in their first year of life are able to form general judgments about categories, that is, to form expectations concerning the properties of as-of-yet unencountered instances of the category (see e.g. Baldwin et al., 1993)." (Leslie 2014, 210)

[146] For a more comprehensive summary of semantic approaches to generics, see Leslie and Lerner 2016.

with accommodating others.

In an attempt to give a more unified account, some theorists have recently tried to account for interesting properties of generics outside of the semantic theory and locate these properties in metaphysics instead. In particular, they presuppose that there is a unifying phenomenon, which they call 'genericity,' that generic sentences instantiate. According to genericists, the truths that generic statements express, thus, become independent generic truths or the truths about *genericity*.

Authors who place the properties of generics in metaphysics instead of semantics also offer a semantics for generics. The main disagreement between the genericity approaches is about what explains what, namely what is explained by the semantics and what by the metaphysics (see Section 5.4.3). For genericists, taking into consideration the correct theory of genericity has, thus, become an important element when assigning truth values to generic statements.

5.3 Generic Judgments and Generic Truths

Speakers often make judgments about the truth of propositions. Consider claims such as:

(1) All people are mortal.

(2) Some people are tall.

(3) Blacks are violent.

It is, however, worth noticing a distinction between: (i) speakers' *assertion* of a certain proposition *as true,* and (ii) a certain proposition *being true*. The former distinction pertains to *pragmatics* while the latter to *semantics.*[147]

When speakers *assert* certain propositions such as (1)–(3), the proposition being expressed can be either *true* or *false.*[148] Whether a certain proposition is *true* or *false* is usually taken to depend on the following conditions: *the phenomenon or the state of affairs in the world*, and *the truth-conditions*[149] *of a particular proposition*. Under the assumption that we have epistemic access to both conditions, the task of assigning *truth values* to propositions should be a relatively simple task.

A mismatch can occur between speakers' judgments and the truth of a proposition when speakers' judgments do not match either one or both of the above two conditions, i.e. *the phenomenon or the state of*

[147] For more about how pragmatics and semantics can come apart in the case of generics, see Chapter 6.

[148] Following classical propositional logic, a proposition is either true or false. For the purposes of this chapter, I do not consider non-classical propositional logic according to which propositions can have truth value other than true or false, or can have an indeterminate (or incomplete, or inscrutable) truth value, or can lack a truth value, or can be both true or false, etc.

[149] Roughly speaking, truth-conditions are conditions under which a given proposition is true, sometimes also taken to be the meaning of a proposition.

affairs in the world, and *the truth-conditions of a particular proposition*. Here are some illustrations.

First. Consider a case of a mismatch between speakers' judgments and *the phenomenon or the state of affairs in the world*. For example, imagine *asserting* "All tigers have stripes" despite the existence of non-striped tigers. Or imagine *rejecting* "None of the students is in the classroom" while there being a student hiding behind the curtains in the classroom.

Second. Consider a case of a mismatch between speakers' judgments and *the truth-conditions of a particular proposition*. For example, imagine *asserting* "All tigers have stripes" while e.g. being aware of the existence of albino tigers yet, for some reason, not acknowledging the truth-conditions of the proposition in question.[150] Or imagine *rejecting* "A student is in the classroom" while being aware of the student hiding behind the curtains in that very classroom yet, for some reason, not acknowledging the truth-conditions of the proposition in question.[151]

Third. Consider a mismatch between speakers' judgments and both *the phenomenon or the state of affairs in the world* as well as *the truth-conditions of a particular proposition*. For example, imagine *asserting* "All tigers have stripes" while, for some reason, not acknowledging the existence of albino tigers nor contribution to truth-conditions made by "all."[152] Or consider *denying* "None of the students is in the classroom" while, for some reason, not acknowledging the fact that the student is hiding behind the curtains in that very classroom nor the truth-conditions of "none."[153]

The above illustrations demonstrate how a mismatch can occur between *speakers' judgments* and the *truth of a proposition.* However, despite a possible mismatch, semanticists of natural language often rely on speakers' intuitions or judgments about the truth of a proposition when offering truth-conditions for a certain type of propositions. Adopting this methodology has proven to be considerably easier task for types of propositions exemplified by statements (1)–(2) than for the type of propositions exemplified by statement (3) (see also Chapter 6, Section 6.4.5.3).

In particular, when it comes to *universally quantified statements* such as (1) there is a high agreement among the speakers of a natural language about what the semanticists have captured under the logical form of a universal quantifier, usually expressed in natural language by "all," or "every." For instance, speakers assert universal statements that have a form "All things have property P if they have property Q" (e.g. "All people are mortal"), under the assumption that every member of the kind (e.g. "people")

---

[150] E.g. imagine that, contrary to the truth-conditions, one presupposes that not every single tiger has to have stripes in order for this proposition to be true, but only all the tigers one has seen so far.

[151] E.g. imagine that, contrary to the truth-conditions, one presupposes that in order for the sentence to be true one has to count in only the students that one can see.

[152] Under the assumption that "all" presupposes that every single tiger has to have stripes in order for the sentence to be true, and not, for instance, only all the tigers they have seen so far.

[153] Under the assumption that "none" presupposes that none of the students must be in the classroom in order for the sentence to be true, and not, for instance, only being out of one's sight.

has to have the property being predicated (e.g. "being mortal").

Similarly, in the case of *existentially quantified statements* such as (2) there is a high agreement among the speakers of a natural language about what the semanticists have captured under the logical form of an existential quantifier, usually expressed in natural language by "there is," or "some," or "at least one." For instance, speakers assert existential statements that have the form "Some things with property Q also have property P" (e.g. "Some people are tall"), under the assumption that at least one member of the kind (e.g. "people") has to have the property being predicated (e.g. "being tall").

However, in the case of *generic statements* such as (3) offering truth-conditions that would accommodate speakers' judgments has proven to be very hard to track. As outlined in Section 5.2, generics come in different syntactic forms, they do not contain any pronounced element in language that would correspond to the content of the generalization being expressed, and their logical structure is not easy to come by.

Given the above, I want to bring out two points about how types of propositions exemplified by *overtly quantified statements* such as (1)–(2) differ from a type of proposition exemplified by *generic statements* such as (3).

Point One. There are different proposals[154] about the truth-conditions of quantified statements such as (1) and (2). However, in the case of quantified statements such as (1) and (2) speakers' judgments are more consistent across the board and, thus, their contribution towards offering different competitive *unified* semantic theories of quantification is more consistent across the board. This is, however, not so in the case of generic statements such as (3) where speakers' judgments seem to point to *disunified* semantics for the same type of proposition, i.e. a generic proposition, and where some semantic theories for generics work better for a certain subclass of generic sentences whereas some other semantic theories work better for some other subclass. In other words, there seems to be a kind of systematicity across different sentences in the case of the types of propositions exemplified by statements such as (1) and (2) that does not seem to exist in the case of the type of a proposition exemplified by statements such as (3).

Point Two. While the advocates of realism about generic truths or facts take generic statements to refer to an independent metaphysical phenomenon of *genericity*, in the case of universal or existential statements there are usually[155] no such advocates that would presuppose realism about universal or existential truths or facts. In other words, when giving semantics for universal and existential

---

[154] For example, in natural language, it is often considered that when uttering sentences such as (1) and (2), what we communicate is restricted by a domain in which the sentence is uttered. There are two dominant views on quantifier domain restriction, namely the so-called "semantic" approach to quantifier domain restriction (see Recanati 1989, 2004; Stanley and Szábo 2000; Neale 2000), and the so-called "pragmatic" approach to quantifier domain restriction (see Bach 1994, 2000; Cappelen and Lepore 2001, 2002, 2005).

[155] Generalized quantification theory might come close to this (see Westerståhl 2016). Thanks to Sam Roberts for pointing this out.

quantifiers it is not presupposed that universal and existential statements refer to some independent metaphysical phenomenon of universal and existential truths or facts in the same way realists about generic truths presuppose the existence of an independent metaphysical phenomenon of genericity.

In the rest of the chapter, I take on board genericists' assumption about the realism of generic truths in order to show a dilemma for the realist theories of genericity. For the purposes of the dilemma, I assume that generic truths are more like evaluative truths rather than universal and existential truths, and I show that, contrary to what genericists believe, we do not have a good reason to believe that generic judgments that speakers make refer to independent generic truths.

5.4 Establishing Analogy

This chapter posits a Darwinian dilemma for realist theories of genericity based on a direct analogy with Street's (2006) Darwinian dilemma for realist theories of value. In her paper called "A Darwinian Dilemma for Realist Theories of Value" Sharon Street (2006) argues against *realist theories of value.* In a nutshell, Street's Darwinian dilemma shows that *realist theories of value* are not able to accommodate a hypothesis that human values are deeply influenced by evolutionary forces. In this chapter, I consider the prospects of applying an analogous Darwinian dilemma for *realist theories of genericity*. In order to establish the analogy, in the next few sections, I identify and introduce some structural analogies between the two debunking arguments.

5.4.1 Preliminaries

In this section, I, first, offer some *motivation* for the analogy (Section 5.4.1.1). Next, I offer some working definitions of preliminary assumptions for the *realist theories* (Section 5.4.1.2), *truths/facts* (Section 5.4.1.3), *judgments* (Section 5.4.1.4), *the relation between judgments and evolutionary forces* (Section 5.4.1.5), that both Street and I need for our Darwinian dilemmas about value and genericity, respectively. Finally, I adopt the argumentative structure of Street's Darwinian *dilemma* and apply it to generics (Section 5.4.1.6).

5.4.1.1 Motivation

One possible venue into applying a direct analogy with Street's dilemma could come from accepting a hypothesis that a certain relation holds between evaluative and generic truths, namely that generic truths are in some way *alike to* evaluative truths. Let us consider four different candidate hypotheses that could capture this relation:

(i) Evaluative truths are a subset of generic truths.

(ii) Some generic truths are equal to some evaluative generics.

(iii) Generic truths are a subset of evaluative truths.

(iv) All generic truths are evaluative truths.

When it comes to their plausibility, it is worth noticing that some connections have recently been made towards hypotheses (i) and (ii). In particular, Thakral (manuscript) argues that moral principles are best understood as generics, which can be seen along the lines of hypothesis (i). It has also been argued that *ceteris paribus* laws of special sciences are plausibly expressed by generics as moral generalizations (see Liebesman 2011, Nickel 2010). Furthermore, certain normative generics[156] (see Leslie and Lerner 2016, Section 5.1) can arguably be understood as an example of the option (ii). If hypotheses (i) or (ii) hold, that would mean that at least some generic truths are susceptible to a Darwinian dilemma. On the other hand, the easier way to extend a Darwinian dilemma for realist theories of value to the realist theories of genericity would be to accept hypotheses (iii) and (iv). I am, however, not aware of any proposals that would argue for such a relation. Thus, even though there are some possible connections to be made between the evaluative truths and generic truths, in this chapter, I do not presuppose that they are necessary for establishing my analogy with Street's (2006) debunking argument. Instead, my claim is weaker, I take them as being strongly motivational for drawing the structural analogy. I rely on the strong analogy between the realist theories about value and the realist theories about genericity on the one hand, and on the the other hand, on the strong analogy between the assumption that natural selection has influenced our evaluative judgments and the assumption that natural selection has influenced our generic judgments.

5.4.1.2 Realist Theories

Realist Theories of Value: claim that there are "at least some evaluative facts or truths that hold independently[157] of all our evaluative attitudes" (Street 2006, 110).

Realist Theories of Genericity: claim that there are at least some generic facts or truths that hold independently of all our generic attitudes.

---

[156] "In addition to fostering essentialist beliefs, generics are also often used to communicate normative statements about social groups, e.g., "boys don't cry", "a woman puts family before career", "friends don't let friends drive drunk". Generics such as these are not intended as descriptions of the facts on the ground, but rather carry a particular normative force; "friends don't let friends drive drunk" was not introduced as a banal observation about what is in fact the case, but rather was introduced precisely to address the fact that, all too often, people were allowing their friends to drive under the influence. In the context of gender, these normative generics can serve as a concise (and child-friendly) means of communicating and sustaining problematic social norms (Leslie 2015b)." (Leslie and Lerner 2016, Section 5.1)

[157] "Antirealists can therefore agree with realists that the truth of a given evaluative judgement holds independently of whether one makes that particular judgement. Where antirealists part ways with realists is in denying that there are evaluative truths which hold independently of the whole set of evaluative judgements we make or might make upon reflection, or independently of the whole set of other evaluative attitudes we hold or might hold upon reflection." (Street 2006, 111)

5.4.1.3 Truths

Evaluative Truths: "facts or truths of the form that X is a normative reason to Y, that one should or ought to X, that X is good, valuable, or worthwhile, that X is morally right or wrong, and so on" (Street 2006, 110).

Generic Truths:[158] facts or truths that are a result of a generic predication[159] of property P to a given kind K.

Furthermore, Street argues that realist theories of value take the evaluative truths to be *stance-independent*[160] from our evaluative attitudes (see Street 2006, 111). For example: "According to *realists [about value]*, the *[evaluative] truth* that *Hitler was morally depraved* holds independently of any stance that we (or Hitler) might take toward that truth, whether now or upon reflection" (Street 2006, 111; italics mine).

A similar stance-independence can be adopted for generics, namely stance-independence of generic truths from our generic judgments can also be applicable. For example, one could claim that, according to realists about genericity, generic truth (if any) that e.g. Muslims are terrorists holds independently of any stance that we (or Muslims) might take toward that truth, whether now or upon reflection.

5.4.1.4 Judgments

Evaluative Judgments: "include states such as desires, attitudes of approval and disapproval, unreflective evaluative tendencies such as the tendency to experience X as counting in favor of or demanding Y, and consciously or unconsciously held evaluative judgements, such as judgements about what is a reason for what, about what one should or ought to do, about what is good, valuable, or worthwhile, about what is morally right or wrong, and so on" (Street 2006, 110).

Generic Judgments: include states such as desires, attitudes of approval and disapproval, unreflective generic tendencies such as the tendency to experience *property P being generically predicated to a given kind K*, and consciously or unconsciously held generic judgments that *property P*

---

[158] See Nickel 2016, 197, for mapping *characteristicness* to *normality*: For a given kind K and set of explanatory strategies S, and for each characteristic property P of Ks relative to S: the property of participating in the mechanism associated with K, S, and P is a way of being a P-normal K relative to S.

[159] How exactly "generic predication" is spelled out depends on a particular theory of genericity that one endorses (see Section 5.3.3).

[160] Street (2006) adopts Shafer-Landau's (2003) notion of stance-independence based on the terminology from Ronald Milo.

is being generically predicated to a given kind K.

5.4.1.5 Judgments and Evolutionary Forces

Relation between Evaluative Judgments and Evolutionary Forces: "the content of human evaluative judgements has been tremendously influenced—indirectly influenced, in the way I have indicated, but nevertheless tremendously influenced—by the forces of natural selection, such that our system of evaluative judgements is saturated with evolutionary influence" (Street 2006, 121).

Relation between Generic Judgments and Evolutionary Forces: the content of human generic judgments has been tremendously influenced—indirectly influenced, but nevertheless tremendously influenced—by the forces of natural selection, such that our system of generic judgments is saturated with evolutionary influence.

5.4.1.6 Darwinian Dilemma

Darwinian Dilemma for Realist Theories of Value: "Evolutionary forces have played a tremendous role in shaping the content of human *evaluative attitudes*. The challenge for *realist theories of value* is to explain the relation be tween these evolutionary influences on our *evaluative attitudes*, on the one hand, and the independent *evaluative truths* that realism posits, on the other. Realism, I argue, can give no satisfactory account of this relation. On the one hand, the realist may claim that there is no relation between evolutionary influences on our *evaluative attitudes* and independent *evaluative truths*. But this claim leads to the implausible skeptical result that most of our *evaluative judgements* are off track due to the distorting pressure of Darwinian forces. The realist's other option is to claim that there is a relation between evolutionary influences and independent *evaluative truths*, namely that natural selection favored ancestors who were able to grasp those truths. But this account, I argue, is unacceptable on scientific grounds. Either way, then, *realist theories of value* prove unable to accommodate the fact that Darwinian forces have deeply influenced the content of *human values*" (Street 2006, 109; italics mine).

Darwinian Dilemma for Realist Theories of Genericity: Evolutionary forces have played a tremendous role in shaping the content of human *generic attitudes*. The challenge for *realist theories of genericity* is to explain the relation between these evolutionary influences on our *generic attitudes*, on the one hand, and the independent *generic truths* that realism posits, on the other. Realism, I will argue, can give no satisfactory account of this relation. On the one hand, the realist may claim that there is no relation between evolutionary influences on our *generic attitudes* and independent *generic truths*. But this claim leads to the implausible skeptical result that most of our *generic judgments* are off track due to the distorting pressure of Darwinian forces. The realist's other option is to claim that there is a relation between evolutionary influences and independent *generic truths*, namely that natural selection favored ancestors who were able to grasp those truths. But this account, I will argue, is unacceptable on scientific grounds. Either way, then, *realist theories of genericity* prove unable to accommodate the fact that Darwinian forces have deeply influenced the content of *human*

genericity.

5.4.2 Two Caveats

In this section, I briefly consider two caveats for a Darwinian dilemma. The first one is connected with the hypothetical nature of the argument (Section 5.4.2.1) and the second one is connected with the limited strength of the argument (Section 5.4.2.2).

5.4.2.1. The Hypothetical Nature

The argument I offer falls prey to a similar caveat that Gibbard (1990) and Street (2006) also recognize, namely my argument presupposes the same conditional form as theirs: "… if the psychological facts are roughly as I speculate, here is what might be said philosophically." Gibbard (1990, 30); or "… if the evolutionary facts are roughly as I speculate, here is what might be said philosophically" (Street 2006, 112).

However, as well as Gibbard (1990) and Street (2006), I also try to rest my arguments on the "least controversial, most well-founded evolutionary speculations possible" Street (2006, 112). Furthermore, my theorizing too falls within evolutionary psychology,[161] a subfield of evolutionary biology. According to evolutionary psychology there is an analogy between cognitive and physical traits. Here is the connection Street points out:

> According to this subfield, human cognitive traits are (in some sense) just as susceptible to Darwinian explanation as human physical traits are (in some cases). For example, a cognitive trait such as the widespread human tendency to value the survival of one's offspring may, according to evolutionary psychology, be just as susceptible to evolutionary explanation as physical traits such as our bipedalism or our having opposable thumbs. (Street 2006, 113)

I believe that a similar thing can be said about our cognitive traits of making certain generic judgments, namely that our judgments such as "Sharks eat swimmers," "Ducks lay eggs," "Blacks are violent," "Women are submissive" are susceptible to evolutionary explanations as well. I develop this connection further in Section 5.4.4.

Furthermore, as Street (2006, 113) points out, there are many pitfalls that such theorizing needs to avoid "the most important of which is the mistake of assuming that every observable trait (whether cognitive or physical) is an adaptation resulting from natural selection, as opposed to the result of any number of other complex (non-selective or only partially selective) processes that could have produced it" Street (2006, 113). However, I believe that despite a possible skepticism about details of evolutionary picture that is offered in this chapter, evolutionary psychology can offer us details that are certain

[161] For introductions to the field of evolutionary psychology, see Barkow et al. 1992 and Buss 1999.

enough to suffice for us looking into its philosophical implications.

5.4.2.2 The Limits

Another caveat I want to bring up is the limited strength of debunking[162] arguments against their targets. On the one hand, there are different types of debunking[163] arguments in the literature. According to Königs (see 2018, 384), we can distinguish between two main types of debunking arguments: one that operates through regular undercutting defeat, namely by undercutting a belief or doctrine by exposing its causal origins, and another one that relies on higher-order evidence. As he puts it:

> The primary argument removes the evidential force of deontological intuitions, which had been thought to support deontology. The secondary argument, by contrast, relies on defeat based on what has come to be called higher-order evidence. Higher-order evidence is "evidence about the character of [the first-order evidence] itself, or about subjects' capacities and dispositions for responding rationally to [the first-order evidence]" (Kelly, 2014). Higher-order defeat occurs when the higher-order evidence suggests that the first-order evidence does not support the proposition in the way it had been taken to support it. Higher-order defeat is thus similar to ordinary undercutting defeat in that the latter, too, severs the evidential relation between a piece or body of evidence and the proposition it is thought to support. However, higher-order defeat is in an important respect different in that it implies that one's assessment of the first-order evidence was flawed to being with. (Königs 2018, 389)

Königs (2018) then raises interesting objections to the second type of debunking argument by claiming that the second type is objectionably sloppy and inadmissible in academic discussion. Nevertheless, these objections do not affect my debunking argument since my argument falls into the first type.

However, even though the type of a debunking argument I rely on is of the first type or the undercutting one, it is worth stressing that, even if successful, the scope of such a debunking argument is limited. In particular, debunking arguments about value do not establish that deontology is false. Instead they can "merely establish that there is no positive reason to believe that deontology is correct" (see Königs 2018, 389). This argument will not rule out that our moral judgments are true (assuming the truth of realism about value) neither it will rule out that there are moral facts (see Street 2006, Joyce 2006, 2013). Similarly, if successful, a debunking argument about genericity will not establish that realism about genericity is false but rather that there is no positive reason to believe that realism about genericity is correct. Furthermore, a debunking argument about genericity will also not rule out that our generic judgments are true (assuming the truth of realism about genericity) neither will it rule out that

---

[162] The logic behind debunking arguments is "that of undermining a belief or doctrine by exposing its causal origins" (Königs 2018, 384).

[163] For other debunking arguments in moral philosophy, see, for example, Cohen 2000; Joyce 2013; Huemer 2013; Morton 2016; and Street 2006. For the debates in philosophy of religion and metaphysics, see, for example, Barrett 2007; Mason 2010; Thurow 2013; and Korman 2014, respectively. Cf. also Königs 2018, 396.

there are generic facts.

5.4.3 The Target: Realist Theories of Genericity

In this section, I introduce and further explain what I take to be the target of my Darwinian dilemma. In particular, first, I offer some motivation for developing a realist theory of genericity (Section 5.4.3.1). Next, I give a working definition and an example of a realism about genericity (Section 5.4.3.2). After that, I distinguish between two kinds of realism about genericity (Section 5.4.3.3). Finally, I briefly glance over some current proposals of realist theories of genericity (Section 5.4.3.4).

5.4.3.1 Motivation

We can distinguish between the two approaches in theorizing about generics: the view that takes the semantic theory to be a theory of genericity—call it *The Standard View,* and the view that offers a separate, metaphysical theory of genericity—call it *Realism about Genericity*.

The Standard View treats "the semantic theory as ipso facto a theory of genericity. All the interesting properties of the intuitive truth-conditions of generics are semantically relevant and to be encoded in their semantics" (Sterken 2016). On the other hand, the advocates of *Realism about Genericity* offer a separate metaphysical theory of genericity presupposing that there exist mind-independent generic facts or truths. As Sterken (2016) puts it:

> Some theorists, faced with the seemingly impossible task of coming up with an adequate semantic theory of generic sentences, eschew semantic theory entirely and locate the interesting properties of generics in the metaphysics of what we might call, genericity. Such a metaphysical theories allowed to be unsystematic in a way that the semantic theory cannot be, allowing such theorists to place the philosophically interesting phenomenon that seems to underlie generic meaning, outside of the semantic theory altogether. (Sterken 2016)

For instance, as Sterken (2016) summarizes it, Nickel's (2016) motivation for giving a metaphysical theory of genericity can be seen as the following:

> In many cases of theorizing about the meaning of particular expressions (for example, theorizing about *the* or *spy*), there is no philosophically interesting underlying metaphysics worthy of theorizing about. There is, for example, no interesting field studying the metaphysics of definiteness, or the metaphysics of *spyhood*. In other cases, there is a rich, independent metaphysics associated with the phenomena that our thought and talk is aiming to capture—for example, modals and modality. Nickel invites us to understand generics and genericity by analogy with modals and modality. (Sterken 2016)

It is worth recalling that the development of realist theories of genericity has been motivated by the difficulty of coming up with a unified semantic theory of generics (see Section 5.2). Many theorists, thus, take the phenomenon of genericity and generic sentences to be in a close relation. For instance, Sterken (2015, 2) argues that the advocates of genericity see the central value of postulating the

phenomenon of genericity in providing a simple and unified semantics for generic sentences, regardless of the variety of ways in which the truth-conditions of generics seem to vary. As she puts it: "The basic theoretical starting point of nearly all theorists has been to assume that there is a unifying phenomenon called genericity, which generics, in general, instantiate: *Assumption of Unity:* There is a unified phenomenon of genericity that generic sentences, in general, instantiate" (Sterken 2015, 2). Furthermore, Sterken (see Sterken 2015, 2) argues that the underlying idea of those who endorse genericity is finding out what the phenomenon of genericity amounts to because they hope this will lead us to a semantics for generic sentences that is unified and simple. Thus, their underlying assumption is that the complexity connected to generics belongs to the phenomenon of genericity rather than the semantics of generics.

5.4.3.2 Working Definition and an Example

As introduced in Section 5.4.1, I understand the realist theories of genericity in the following manner:

> *Realist Theories of Genericity:* claim that there are at least some generic facts or truths that hold independently of all our generic attitudes.

Furthermore, I take it that the realist theories of genericity share the following three features:

(i) they place genericity outside of semantic theory;
(ii) they offer a metaphysical theory of genericity;
(iii) they take genericity to be a mind-independent phenomenon.

For example, Nickel (2016) offers an entirely separate theory of genericity from a semantic theory of generics, and argues for their independency. However, these two theories will be closely connected in the sense that genericity can be manifested through the meaning of generics and that we get to "know" metaphysical phenomena, i.e. genericity, through the linguistic phenomena, i.e. generic sentences. In more detail, here is how Nickel separates the metaphysical phenomenon of genericity from language:

> This theory of characteristicness [genericity], framed in terms of mechanisms, is a basically metaphysical theory. (…) We express genericity in language using tools—expressions and constructions—that also see employment in non-generic modes of speech. Generics can be logically complex, they can be plural, and they contain predicates that are semantically complex in ways that are independent of genericity. (Nickel 2016, 7–9)

Furthermore, apart from placing genericity outside of semantic theory and offering a metaphysical

theory of genericity, Nickel (2016) also takes genericity to be a mind-independent phenomenon:

> My theory thus tries to capture, within a uniform theoretical framework, the obvious fact that some generics are fairly objective, while others are not. Yet I want to emphasize that on this theory, not anything goes. There is a component of human convention and interest to the interpretation of generics if I'm right, but that is only one component. The other component is completely objective: it is the existence of a mechanism. Without a mechanism connecting at least some members of the kind to the property predicated in the generic, the generic cannot be true. So according to my theory, generics that reflect the interests and (attempted) explanatory strategies of astrology one and all fail to be true, simply because there is no mechanism by which the stars come to influence our lives. *Explanatory and investigative interests only serve to separate the suitable from the unsuitable mechanisms. They cannot bring mechanisms into existence.* (Nickel 2016, 7; italics mine)

Both the mind-independence and the independence of genericity from the linguistic phenomena and, thus, also (at least indirectly) from our generic judgments, makes realists about genericity of this kind an ideal target of the Darwinian dilemma.

5.4.3.3 Two Kinds of Realism[164]

Street (2006, 111–112) distinguishes between the two camps of realism about value: (i) *non-naturalist value realism,* and (ii) *value naturalism*. Both camps share the view about the stance-independence between evaluative truths and evaluative attitudes. However, the two camps differ in "how they construe the nature of these facts or truths." (Street 2006, 111)

Non-naturalist Value Realism: evaluative truths are not reducible to natural facts and are causally inert, they are rather "irreducibly normative facts or truths" (Street 2006, 2012).

This is the kind of naturalism that is a primarily target of Street's argument.

Value Naturalism: a position which "holds that evaluative facts are identical with or construed by (certain) natural facts, and that evaluative facts are the kinds of things that play a role in causal explanations" (Street 2006, 112).

Even though the latter kind of naturalism is not her primary target she believes that at least certain forms of this kind of naturalism are also targeted by her argument (see Street 20016, 112, and Section 7).

Similarly, I also suggest to distinguish between two camps of naturalism about genericity.

---

[164] Quasi-realism is another candidate of realism that Street distinguishes but she is explicit about focusing only on "non-quasi" realism (see Street 2006, 157). For the purposes of this chapter, I also do not consider the quasi-realism.

Non-naturalist Genericity Realism: generic truths are not reducible to natural facts and are causally inert, they are rather irreducibly generic facts or truths.

Genericity Naturalism: a position which holds that generic facts are identical with or construed by (certain) natural facts, and that generic facts are the kinds of things that play a role in causal explanations.

5.4.3.4 Some Current Proposals

The target of Darwinian dilemma I offer is not any particular theory of genericity, but realist theories in general. Given the scope of my target, it will, thus, suffice to briefly mention some candidates of realism about genericity. In doing so, I draw on Sterken (2015, 2016) who identifies some recent proposals that do not take the semantic theory to be a theory of genericity but, instead, offer a metaphysical theory of genericity. Here are, thus, some rough ways of how to conceive what the metaphysics of genericity is:

Generalizations involving a notion of normality: e.g. Nickel (2008, 2016, 2017),[165] and Asher and Pelletier (2012);

Instantiations of inheritance relations between kinds and their instances: e.g. Liebesman (2011);[166]

Indexicality and context-sensitivity: e.g. Sterken (2015);[167]

Probabilistic (majority) generalizations: e.g. Cohen (1996);

Fundamentally mental phenomenon: e.g. Leslie (2007, 2008).[168]

5.4.4 Opening Premise: Generic Judgments are Influenced by Evolutionary Forces

In this section, I introduce and argue for an opening premise of a Darwinian dilemma for realist theories of genericity, analogous to Street's (2006) opening premise of a Darwinian dilemma for realist theories of value. In particular, drawing on Street's opening premise as well as on the evidence from evolutionary theory and evolutionary and cognitive psychology, I argue that there is a connection

---

[165] "Nickel classifies genericity as requiring a "particularly close or intimate connection between a kind and a property, one that does not obviously coincide with either statistical notions (all, most, many, some), nor does it coincide with well-established modal notions (necessity or essence)" (2017, p. 440). He justifies the claim given that generics have modal import that requires more than statistical correlations and are most acceptable when they involve well-established kinds." (Ritchie 2019, 35)

[166] Liebesman (2011) takes genericity to be about how kinds can inherit properties from their members. How this happens does not have to be in the same way in every case. As Sterken (2016) points out, this way of thinking traces back to the early Greg N. Carlson (1977).

[167] Sterken positions herself among those who offer a theory of genericity and claims that she offers "a metasemantic account of the theoretically interesting properties of generic meaning." (Sterken 2016)

[168] Leslie talks about the "worldly truth-makers" of generics (see Leslie 2007, 386) and takes generics to be instantiations of our cognitively primitive generalizations (see Leslie 2007, 2008).

between generic judgments and evolution, namely I argue that evolutionary forces have played a role in shaping the content of human generic attitudes.

In Section 5.4.4.1, I present Street's and my working opening premise of a Darwinian dilemma. In Section 5.4.4.2, I consider two worries Street addresses and show how they apply to my opening premise. In Section 5.4.4.3, I present a refinement of Street's opening premise and apply this refinement to my opening premise.

5.4.4.1 The Working Opening Premise

In this section, I want to utilize certain general claims that Street (2006) makes about evaluative judgments that, I believe, also apply to generic judgments. In addition to that, I offer some support from cognitive science and developmental psychology in order to support some of the more specific claims connected with our basic generic tendencies.

Street's (see 2006, 113–121) opening premise, in its first approximation, says that natural selection has played a tremendous role in influencing the content of our evaluative judgments. She, however, does not want to discount other possible influences, both evolutionary and non-evolutionary. As she puts it:

> No doubt there have been numerous other influences: some of them were perhaps evolutionary factors other than natural selection—for example, genetic drift; and many other forces were not evolutionary at all but rather social, cultural, historical, or of some other kind. And then there is the crucial and *sui generis* influence of rational reflection that must also be taken into account. I am discounting none of these influences. My claim is simply that one enormous factor in shaping the content of human values has even the forces of natural selection, such that our system of evaluative judgements is thoroughly saturated with evolutionary influence. (Street 2006, 113–114)

Similarly, I want to argue that apart from non-evolutionary influences such as, for example, rational reflection and factors like social, cultural, historical ones, there have been evolutionary influences that shaped our generic judgments. I, thus, believe that we can utilize a Darwinian explanation, in terms of the benefits and costs of natural selection for the reproductive success in order to argue for the influence of evolutionary forces on our generic judgments. For instance, we can assume that different generic tendencies we have can contribute to creating different chances of our survival and reproduction. Given that, we can then presuppose that there has been a certain selective pressure on the content of our generic judgments or their "proto"[169] versions throughout our evolutionary history. In other words, on the one hand, there has been selective pressure on making generic judgments which tended to promote our survival and reproductive success. On the other hand, there has been selective pressure on refraining from making those generic statements which tended to decrease our

[169] See Street (2006, 114–115) for the introduction of the idea of the "proto" version of our evaluative judgments.

survival and reproductive success. As a result, natural selection favored ancestors who were able to grasp those judgments which tended to increase our survival.

Furthermore, Street appeals to the patterns of evaluative judgments people tend to make and claims that despite their unlimited diversity "there are deep and striking patterns, across both time and cultures, in many of the most basic evaluative judgments that human beings tend to make" (2006, 115). Moreover, she argues that evolutionary biology can explain these patterns because they tended to promote survival and reproduction as opposed to the alternative ones. However, rough explanations[170] for the promotion of survival and reproduction can vary depending on the evaluative tendency in question and can be more or less complicated, and can appeal to different parts of evolutionary theory.

Street (2006, 115–117) offers two pieces of evidence in support of the fact that the content of our evaluative judgments has been greatly influenced by natural selection. First, she considers judgments that are opposite to the ones that we accept, such as "the fact that something would promote one's survival is a reason against it" or "we have greater obligations to help complete strangers than we do to help our own children," etc. If, other things being equal, these were the judgments widely held across both time and cultures, and they would increase the reproductive and survival success of their promotors, then that would be the counter-evidence for claims such as "the fact that something would promote one's survival is a reason in favor of it" or "we have greater obligations to help our own children than we do to help complete strangers" being the ones that have been greatly influenced by natural selection.

Analogously, I believe that we can find similar tendencies to promote the success of survival by accepting certain generic statements (as opposed to their negation). For example, consider generics such as:

a. Candy is bad for your teeth.

b. A raven is black.

c. The tiger has stripes.

d. Ticks carry Lyme disease.

e. Sharks eat swimmers.

f. Lions have manes.

g. Muslims are terrorists.

h. Mosquitos carry the West Nile virus.

---

[170] It is also important to notice that: "The explanandum is *not* particular attitudes held by particular individuals—for example, your or my or George W. Bush's judgement that the fact that something would help a family member is reason to do it. Such individual-level facts are not appropriate objects of evolutionary explanation. What are appropriate objects of evolutionary explanation are population-level facts about patterns of variation in a given trait across a population—and the widespread presence of certain basic evaluative tendencies in the human population are such objects" (Street 2006, 158).

Psychologically oriented theorists have argued that generic generalizations are sensitive to the content-based factors or the nature of the property being generalized. Such factors include *whether the property being predicated is dangerous* (see Cimpian, Brandone and Gelman 2010). For example, consider generics such as "Mosquitoes carry West Nile virus," "Sharks attack swimmers" or "Ticks carry Lyme disease." Another content-based factor includes *whether the property being predicated is characteristic of the kind* (Prasada and Dillingham 2006, 2009; Leslie 2007, 2008; Gelman and Bloom 2007; Cimpian, Brandone and Gelman 2010; Cimpian, Gelman and Brandone 2010; Prasada et al. 2013). For example, consider generics about distinctive physical features such as "Lions have manes" or generics about methods of reproduction for animal kinds such as "Ducks lay eggs" or generics about functions for artifact kinds such as "Orange-Crusher-2000s crush oranges."

It is interesting to notice that, in such cases, generics can be accepted at a rather low prevalence (see Leslie 2008). One plausible explanation can be that we make these generic judgments because that has been evolutionary beneficial for us to do. For instance, Leslie (2017) recognizes the connection between accepting a certain subclass of generics she dubs *striking property generics* and its evolutionary benefits:

> Since we are working under the hypothesis that generics give voice to psychologically fundamental generalizations, this observation implies that our basic way of dealing with dangerous or harmful information involves rapidly generalizing this information to the salient kind or category. We do not wait around to see what percentage of tigers eat people before drawing a general conclusion—even a single instance may be enough for us to conclude that tigers eat people. It is not hard to imagine the evolutionary benefits of such a disposition, since the costs of under-generalizing such information are potentially huge. Our ancestors were far better off jumping to conclusions, as it were, than taking the time to judiciously determine the precise likelihood of their being eaten. (Leslie 2017, 396)

The second piece of evidence Street (2006, 117) offers rests on her observation that other animals that are closely related to us (e.g. chimpanzees) also have basic evaluative tendencies (e.g. recognizing *counting in the favor* or when *other animal has helped them*) that are similar to our widely held evaluative judgments. In this continuity she sees support for the claim that evolutionary forces have played a great role in shaping the content of our evaluative judgments.

Analogously, I believe that it is plausible to think that there are something like basic generic tendencies in other animals too, i.e. that we could observe certain behaviors of some animals as their generic tendencies, e.g. when certain animals *generally* avoid their predators, or when some insects and birds *generally* form flocks in order to stave off a predator. Even though more work is needed to subject these claims to scientific testing, I believe there is a basic plausibility in a claim that if there is a continuity between us and our close animal relatives with respect to our basic generic tendencies, then

evolutionary forces have played a role in shaping our evaluative judgments.

In fact, evidence from developmental psychology shows that generics are acquired early in development, by approximately 30 months of age (see Gelman and Raman 2003; Gelman, Goetz, Sarnecka and Flukes 2008; Graham, Nayer and Gelman 2010). For example, some evidence is pointing in the direction that infants and kids in their first year of life show such basic tendencies with respect to kinds and categories (see Baldwin, Markman, and Melartin 1993), as well as that by the mid-preschool years, children hold richly structured and detailed beliefs about the general properties of members of kinds (see Gelman 2003). Furthermore, preschool children have tendencies to evaluate universal or existential statements as if they were generic (see Hollander, Gelman and Star 2002; Tardif, Gelman, Fu and Zhu 2010; Leslie and Gelman 2012; Brandone, Gelman and Hedglen 2015). On the basis of the above mentioned evidence, several researchers have, thus, proposed that generic sentences articulate cognitively fundamental, default generalizations, and that quantified statements, in contrast, articulate cognitively more sophisticated generalizations[171] (e.g. Leslie 2007, 2008, 2012, 2014; Gelman 2010; Cimpian and Erickson 2012; see also Section 5.2.3). More recently, Csibra and Shamsudheen (2015, 705) also argued that "if an object can act as an ad hoc symbol of its kind in communication, a demonstration on this object amounts to a nonverbal generic statement about its kind." This research, I believe, counts towards the support of the hypothesis about the evolution of "proto" forms of generic tendencies into generic judgment and generalizations, more generally.

5.4.4.2 Two Worries

Street considers two worries against her working assumption that evolutionary forces have played a significant role in our evaluative judgments. Since these worries may extend to my analogous assumption that evolutionary forces have played a significant role in our generics judgments, I address them jointly.

The First Worry: Capacity to make evaluative/generic judgments does not precede the selection of their content.

The idea behind this worry is that the opening premise suggests that in the course of human history our evaluative/generic judgments came first and then came the tendencies to make a certain evaluative/generic judgment rather than another. However, evolution happens neither in these two

[171] "These general judgments formed by the preverbal infant—we might naturally term these generalizations cognitively fundamental generalizations— represent our most basic form of generalizing, one which we exploit from our earliest days. When we grow up to learn our native language, it is natural to suppose that this language will provide us with some means of giving voice to these generalizations. The hypothesis of Leslie (2007, 2008) is that generic sentences are precisely this means of giving voice to our cognitively fundamental generalizations." (Leslie 2014, 210)

stages nor in that order. As Street points out:

> Behavioral and motivational tendencies in the direction of at least some of the pairings of circumstance and response (…) presumably arose and became entrenched in our ancestors long before the rise of any capacity for *full-fledged evaluative judgement*—where I am understanding the capacity for full-fledged evaluative judgement to involve not only an unreflective capacity (…), but also a reflective, linguistically—infused capacity to judge (...), and to step back from such judgements and call them into question. (Street 2006, 118)

Street responds to this worry by embracing the order of explanation the objector gives, namely by claiming that our full-fledged evaluative judgments are "relatively late evolutionary add-on, superimposed on top of much more basic behavioral and motivational tendencies" (Street 2006, 118). I believe that a similar answer can be offered for generic judgments, namely one could claim that first we had some basic behavioral and motivational generic tendencies which then gradually developed into a more complex capacities of generic judgments that are reflective and linguistically-infused.

The Second Worry: Natural selection operates on traits that are genetically heritable.

This worry relies on the claim that natural selection operates on traits that are genetically heritable. Yet our *full-fledged evaluative/generic judgments* with a given content such as "one ought to help those who help you" or "mosquitos carry the West Nile virus" do not seem to exhibit such a genetically heritable trait. Instead, any variations in a given population to make such judgments would not be due to genetic differences but would rather be due to factors such as culture or upbringing.

The way Street (2006, 119) addresses this worry is by introducing a plausible hypothesis that our "more basic evaluative tendencies" were genetically heritable traits and that they were "unreflective, non-linguistic, motivational tendency" to experience something as e.g. "counting in favor for" something else. She calls them "'proto' forms of evaluative judgment." And this kind of basic judgment would be due to genetic difference (see Street 2006, 119). Similarly, one could say that there are *more basic generic tendencies* which can be considered as genetically heritable traits. Such basic generic tendencies can be also seen as unreflective, non-linguistic, motivational tendencies to experience something as having a certain property that is applicable to a certain kind. One could call such judgments *'proto' forms of generic judgments*. Finally, one could argue that such basic generic judgments would also be due to genetic difference.

5.4.4.3 Opening Premise Refined

Finally, Street (2006) offers a refined version of her opening premise. She refines her initial assumption by claiming that even though significant, the influence of natural selection on the content of human evaluative judgments is only *indirect*. She claims that natural selection has instead had a *direct*

and tremendous influence on more basic evaluative tendencies which have, in result, strongly influenced the acceptance of certain evaluative judgments.

Moreover, another important point that Street makes is that our *full-fledged evaluative judgments* are not always lined up in content with our *basic evaluative judgments* due to other factors such as the fact that human beings are reflective[172] and "as such are capable of noticing any given evaluative tendency in ourselves, stepping back from it, and deciding on reflection to disavow it and fight against it rather than to endorse the content suggested by it" (Street 2006, 120).

She also considers the possibility of us being able to be very flexible with respect to our evaluative judgments, given the environmental conditions we are in. However, she believes that basic evaluative tendencies will still make us more likely predisposed for making certain evaluative judgments rather than others. As she puts it:

> Indeed, it is likely that we were selected above all else to be extremely flexible when it comes to our evaluative judgements—not locked into any particular set of them but rather able to acquire and adjust them in response to the conditions in which we find ourselves. In suggesting that we possess basic evaluative tendencies, then, I am simply suggesting that when it comes to certain core issues such as our individual survival, the treatment of our offspring, and reciprocal relations with others, there are likely to be strong predispositions in the direction of making some evaluative judgements rather than others (…). (Street 2006, 158)

It is also important to notice that had there been a difference in the general content of our basic evaluative tendencies, there would also have been a difference in the general content of our full-fledged evaluative judgments, and in loosely corresponding ways. Street offers examples of unreflective evaluative tendencies in: *lions* (male lions have unreflective evaluative tendencies "demanded by the circumstance" to kill the offspring that is not his own and female lions "don't hold it against" him when he kills her offspring in such circumstances); *bonobos* (who engage in sexual relations with different partners as "called for" in all kinds of different circumstances); and *social insects* (who value the welfare of the whole community over their own individual welfare). The point Street makes is that had our basic unreflective judgments been similar to theirs, our full-fledged, reflective judgments would have also looked in a way that would loosely reflect those basic ones (see Street, 2006, 120–121).

Finally, Street concludes that the rough picture that presupposes this indirect influence of natural selection on our evaluative judgments is enough to get the Darwinian Dilemma off the ground.

One can apply a fully analogous refinement of the opening premise for generic judgments, namely that our generic judgments have been indirectly yet significantly influenced by the forces of natural selection.

---

[172] Chapter 6 of this dissertation is an attempt of changing some of our *generic tendencie*s by utilizing the methodological tool of philosophical engineering.

5.5 A Darwinian Dilemma for Realist Theories of Genericity

If we assume that the Opening Premise is true, namely that evolutionary forces have indirectly yet significantly influenced our generic judgments, the main question becomes what is, if any, the relation between our generic judgments and the independent generic truths posited by the realist? The realist about genericity could either *resist* the charge or *accept* the charge by either denying or asserting that relation. I discuss the prospects of a denial of this relation (the first horn of the dilemma) in Section 5.5.1, and its assertion (the second horn of the dilemma) in Section 5.5.2.

5.5.1 The First Horn

The first horn of the dilemma focuses on the realist about value/genericity *denying* the relation between evolutionary influences on our *evaluative/generic judgments* and independent *evaluative/generic truths*. In this section, first, I present Street's argument for denying the relation between evaluative truths and evaluative judgments and I apply an analogous argument for denying the relation between generic truths and generic judgments (Section 5.5.1.1). After that, I consider an objection to Street's argument and raise an analogous objection to my argument (Section 5.5.1.2). Finally, I present Street's reply to this objection and apply an analogous reply to the objection to my argument (Section 5.5.1.3).

5.5.1.1 Denying the Relation

Realism about Value

The first horn of a Darwinian dilemma for realist theories of value presupposes the following:

(A) Darwinian Assumption about Evaluative Judgments: Our *evaluative* judgments have been indirectly yet significantly influenced by natural selection.

(B) Realist's Thesis about Value: There are independent *evaluative* truths.

Realist's Denial of the Relation between (A) and (B): Even if our *evaluative* judgments have been indirectly yet significantly influenced by natural selection, there is no relation between our *evaluative* judgments that have been indirectly yet significantly influenced by natural selection and the independent *evaluative* truths.

Street (2006) considers two possible results[173] following from this denial: (i) *evaluative* judgments are true by sheer chance, or (ii) *evaluative* judgments are off track. Street (2006, 121–125) offers the following explanation:

> Of course it's possible that as a matter of sheer chance, some large portion of our evaluative judgements ended up true, due to a happy coincidence between the realist's independent evaluative truths and the evaluative directions in which natural selection tended to push us, but this would require a fluke of luck that's not only extremely unlikely, in view of the huge universe of logically possible evaluative judgements and truths, but also astoundingly convenient to the realist. Barring such a coincidence, the only conclusion remaining is that many or most of our evaluative judgements are off track. This is the far-fetched skeptical result that awaits any realist who takes the route of claiming that there is no relation between evolutionary influences on our evaluative judgements and independent evaluative truths. (Street 2006, 122)

Next, I show how a full analogy with the first horn of Street's argument for realism about value looks for the first horn of realism about genericity.

Realism about Genericity

The first horn of Darwinian dilemma for realist theories of genericity presupposes the following:

(A) Darwinian Assumption about Generic Judgments: Our *generic* judgments have been indirectly yet significantly influenced by natural selection.

(B) Realist's Thesis about Genericity: There are independent *generic* truths.

Realist's Denial of the Relation between (A) and (B): Even if our *generic* judgments have been indirectly yet significantly influenced by natural selection, there is no relation between our *generic* judgments that have been indirectly yet significantly influenced by natural selection and the independent *generic* truths.

Similarly, two possible results follow: (i) *generic* judgments are being true by sheer chance, or (ii) *generic* judgments are off track. As in the case of realism about value, the results of denying the relation in the case of realism about genericity are either implausible or lead to skepticism. In other words, on the one hand, it is highly unlikely that we have somehow, by sheer chance, started making generic judgments that match the independent generic truths. On the other hand, given the high implausibility

[173] Greene (2008) too observes that "it is unlikely that inclinations that evolved as evolutionary by-products correspond to some independent, rationally discoverable moral truth" (Greene 2008, 72).

of the former option, we are left with another option, namely the one that our generic judgments are (mostly) off track, which leaves us with a highly skeptical result that the realist's about genericity must address.

5.5.1.2 Realist's Comeback

Street (2006, 122–123) considers the following reply from the realist about value. Given that we are thinking creatures and do not merely endorse whatever evaluative tendencies were implanted in us by evolutionary forces, the realist could say that Street's objection is ignoring the power of *rational reflection*. Rational reflection has been another kind of tremendous influence on our system of evaluative judgments that is considered to be related to the truth. As she puts it: "Over the course of human history, endless amounts of reflection have gone on and greatly altered the shape of our evaluative judgements. According to the objection at hand, just as a compass and a little steering can correct for the influence of the wind and tides on the course of one's boat, so rational reflection can correct for the influence of selective pressures on our values" (Street 2006, 123).

Given the analogy, a similar objection applies to the first horn of realism about genericity. Namely, the realist about genericity can claim that my argument ignores the power that the influence of rational reflection has on the system of our generic judgments. In other words, the objection goes that we do not just endorse generic tendencies implanted in us by evolutionary forces. Instead, rational reflection has the power of changing the shape of our generic judgments and can correct the influence of selective pressures on our generic judgments (cf. Chapter 6).

So far, we have seen the first horn of the dilemma that focuses on the realist about value/genericity denying the relation between evolutionary influences on our *evaluative/generic judgments* and the independent *evaluative/generic truths*. The objector to such realism may claim that this leads to an implausible skeptical result about our *evaluative/generic judgments* being off track. The realist's comeback consists in claiming that we can get our *evaluative/generic judgments* back on track by utilizing the power of rational reflection.

5.5.1.3 Reply to the Realist

Street (2006, 123–125) offers a response to this potential comeback of the realist. First of all, instead of insisting on perceiving humans "as automatons who simply endorse whatever evaluative tendencies are implanted in us by evolutionary and other forces" (Street 2006, 123), Street admits the influence of rational reflection on our acceptance of evaluative judgments. However, she argues that the fact that we are using rational reflection presupposes that "something must happen—that something must change—when we become conscious of any foreign influence (such as these Darwinian forces) on our evaluative

judgements" (Street 2006, 123). Furthermore, Street argues, the objector relies on an assumption that is false, namely the objector relies on the assumption that our rational reflection is an "uncontaminated tool" which has been helping us to sort the true from the false evaluative judgments. Instead, Street (2006, 124–125) argues, our rational reflection is as contaminated as the rest of our evaluative judgments.[174] Here is the gist of her reply:

> Thus, if the fund of evaluative judgements with which human reflection began was thoroughly contaminated with illegitimate influence—and the objector has offered no reason to doubt this part of the argument—then the tools of rational reflection were equally contaminated, for the latter are always just a subset of the former. It follows that all our reflection over the ages has really just been a process of assessing evaluative judgements that are mostly off the mark in terms of others that are mostly off the mark. And reflection of this kind isn't going to get one any closer to evaluative truth, any more than sorting through contaminated materials with contaminated tools is going to get one closer to purity. (Street 2006, 124)

There is an analogous reply I want to utilize to respond to the above described analogous objection coming from the realist about genericity. Namely, I too want to admit the influence of a rational reflection on our acceptance of generic judgments. However, acknowledging rational reflection presupposes the existence of other kinds of influences on our generic judgments, among others it presupposes the influence of evolutionary forces. Moreover, rational reflection is not an "uncontaminated tool" that helps us sort out true from false generic judgments. Instead, our rational reflection is contaminated too and, thus, our rational reflection is not getting us any closer to the independent generic truth.

In particular, in Chapter 6, I consider recent *rational reflection* in philosophy that has been utilized in the case of the so-called pernicious (social kind) generics such as "Muslims are terrorists" or "Blacks are violent" by looking at different accounts of engineering in philosophy that try to ameliorate the pernicious effects of such generics (see Haslanger 2011; Leslie 2017; Saul 2017; Ritchie 2019), after which I offer an account that utilizes rational reflection and takes *generics judgments* as objects of engineering. It is worth noting that utilizing rational reflection in the way that I propose does not get us closer to generic truths, i.e. I do not claim that engineering generic judgments with help of rational reflection does get us closer to generic truth.

5.5.2 The Second Horn

The second horn of the dilemma focuses on the realist about value/genericity *asserting* the relation between evolutionary influences on our *evaluative/generic judgments* and independent *evaluative/generic truths*. In this section, first, I describe Street's assertion of the relation between evaluative truths and

[174] "The widespread consensus that the method of reflective equilibrium, broadly understood, is our sole means of proceeding in ethics is an acknowledgment of this fact: ultimately, we can test our evaluative judgements only by testing their consistency with our other evaluative judgements, combined of course with judgements about the (non-evaluative) facts." (Street 2006, 124)

evaluative judgments and apply an analogous assertion of the relation between generic truths and generic judgments (Section 5.5.2.1). Next, I introduce two *scientific evolutionary explanations* about why we make certain evaluative judgments that Street considers, namely *the tracking account* and *the adaptive link account*; after which I utilize a direct analogy in order to offer two analogous scientific explanations for why we make certain generic judgments (Section 5.5.2.2). Finally, I summarize Street's results of why *the adaptive link account* fares better on scientific merits than *the tracking account*; and show that similar results apply in the case of generic judgments (Section 5.5.2.3).

5.5.2.1 Asserting the Relation

Realism about Value

The second horn of a Darwinian dilemma for realist theories of value presupposes:

(A) Darwinian Assumption about Evaluative Judgments: Our *evaluative* judgments have been indirectly yet significantly influenced by natural selection.

(B) Realist's Thesis about Value: There are independent *evaluative* truths.

Furthermore, in contrast to the first horn, the second horn presupposes:

Realist's Assertion of the Relation between (A) and (B): There is a relation between our *evaluative* judgments that have been indirectly yet significantly influenced by natural selection and the independent *evaluative* truths.

Realism about Genericity

The second horn of Darwinian dilemma for realist theories of genericity presupposes:

(A) Darwinian Assumption about Generic Judgments: Our *generic* judgments have been indirectly yet significantly influenced by natural selection.

(B) Realist's Thesis about Genericity: There are independent *generic* truths.

However, in contrast to the first horn, the second horn presupposes:

Realist's Assertion of the Relation between (A) and (B): There is a relation between our *generic*

judgments that have been indirectly yet significantly influenced by natural selection and the independent *generic* truths.

5.5.2.2 Explaining the Relation[175]

It is worth noting that asserting the relation seems to be a more natural route for the realist about *value/genericity* to take since we think that a lot of our *evaluative/generic* judgments are true. Furthermore, asserting the relation presupposes that there is an overlap between the content of the independent *evaluative/generic* truths and the content of our *evaluative/generic* judgments indirectly influenced by evolutionary forces. This overlap calls for an explanation from the realist, namely an explanation about what kind of relation there is between the independent *evaluative/generic* truths and selective pressures on our *evaluative/generic* judgments.

Next, I contrast two competing scientific explanations for the assertion of the above mentioned relation, namely the Tracking Relation Account as the realist's scientific explanation and the Adaptive Link Account as a non-realist, alternative scientific explanation. It is worth noticing that they are two evolutionary accounts of why we tend to make some *evaluative/generic* judgments rather than some others.

Realist's Scientific Explanation: The Tracking Relation Account

Street (2006, 125) considers a possible scientific explanation[176] that the realist about value can offer a tracking[177] relation account.

> *The Tracking Relation Account for Value:* "According to this hypothesis, our ability to recognize evaluative truths, like the cheetah's speed and the giraffe's long neck, conferred upon us certain advantages that helped us to flourish and reproduce. Thus, the forces of natural selection that influenced the shape of so many of our evaluative judgments need not and should not be viewed as distorting or illegitimate at all. For the evaluative judgments that it proved most selectively advantageous to make are, in general, precisely those evaluative judgements which are true."

---

[175] In the next three sections, I consider a direct analogy (unless stated otherwise) between the two evolutionary scientific explanations about why we tend to make some *evaluative* judgments rather than some others that Street (2006) considers and the two evolutionary scientific explanations about why we tend to make some *generic* judgments rather than some others.

[176] Street (2006) uses Parfit's points in order to motivate the tracking account. In particular, Parfit claims: "just as cheetahs were selected for their speed, and giraffes for their long necks, the particular feature for which we were selected was our ability to respond to reasons and to rational requirements" (Parfit, in correspondence).

Nozick also makes a similar point by claiming that: "it seems reasonable to assume there has been some evolutionary advantage in acting for (rational) reasons. The capacity to do so, once it appeared, would have been selected for. Organisms able and prone to act for (rational) reasons gained some increased efficiency in leaving great-grand progeny" (Nozick 1981, 337).

[177] Both Street (see 2006, 159) and I borrow this term from Nozick (1981) who uses "tracking" in similar contexts.

(Street 2006, 126)

We can imagine a realist about genericity offering an analogous tracking relation account for genericity.

> *The Tracking Relation Account for Genericity:* According to this hypothesis, our ability to recognize generic truths, like the cheetah's speed and the giraffe's long neck, conferred upon us certain advantages that helped us to flourish and reproduce. Thus, the forces of natural selection that influenced the shape of so many of our generic judgments need not and should not be viewed as distorting or illegitimate at all. For the generic judgments that it proved most selectively advantageous to make are, in general, precisely those generic judgments which are true.

A crucial feature of the tracking relation account for *value/genericity* is that it functions as a *scientific explanation* that explains the presence of certain *evaluative/generic* judgments rather than some others by utilizing the evolutionary explanation, namely that developing the capacity to track *evaluative/generic* truths has proven to be beneficial in terms of survival and reproduction. For example, an advocate of the tracking relation account for genericity may claim that our generic judgments such as "Sharks eat bathers" or "Mosquitos carry the West Nile virus" are true and that being able to track and grasp such truths promoted our reproductive success.

As Street notices (see Street 2006, 159), a non-naturalist about value is forced "to take a stand on certain matters of scientific explanation." One could, thus, object that a non-naturalist would need to explain why humans tend to make some *evaluative/generic* judgments rather than others, as well as how have we developed the capacity to grasp the independent *evaluative/generic* truths. It seems that for the non-naturalist about value/genericity doing so without positing any causal role of *values/genericity* in our scientific explanations is not a viable option, unless the non-naturalist opts for the first horn and denies the relation between our *evaluative/generic* judgments and the independent *evaluative/generic* truths.

It is crucial to notice that, as a scientific explanation, the tracking relation account for v*alue/genericity* has competitors in all other scientific explanations that offer an account of why we tend to make certain *evaluative/generic* judgments rather than others. In particular, Street (2006) considers one such alternative scientific account—the adaptive link account—as another scientific explanation that also considers evolution in order to explain why we tend to make some evaluative judgments rather than others.

Alternative Scientific Explanation: The Adaptive Link Account

Instead of presupposing evaluative truths, the adaptive link account explains evaluative judgments in terms of adaptive links between the circumstances and our response to these circumstances:

> *The Adaptive Link Account for Value:* "tendencies to make certain kinds of evaluative judgements rather than others contributed to our ancestors' reproductive success not because they constituted perceptions of independent evaluative truths, but rather because they forged adaptive links between our ancestors' circumstances and their responses to those circumstances, getting them to act, feel, and believe in ways that turned out to be reproductively advantageous. To elaborate: As a result of natural selection, there are in living organisms all kinds of mechanisms that serve to link an organism's circumstances with its responses in ways that tend to promote survival and reproduction." (Street 2006, 127)

Similarly, one can offer an analogous adaptive link account for genericity.

> *The Adaptive Link Account for Genericity:* tendencies to make certain kinds of generic judgments rather than others contributed to our ancestors' reproductive success not because they constituted perceptions of independent generic truths, but rather because they forged adaptive links between our ancestors' circumstances and their responses to those circumstances, getting them to act, feel, and believe in ways that turned out to be reproductively advantageous. To elaborate: As a result of natural selection, there are in living organisms all kinds of mechanisms that serve to link an organism's circumstances with its responses in ways that tend to promote survival and reproduction.

It is crucial to stress that, according to *the adaptive link account*, evaluative/generic judgments contributed to our ancestors' reproductive success not because they were true or false, as opposed to the *tracking account* according to which our ancestors' reproductive success rests on evaluative/generic judgments that are true. Some of the examples of the adaptive link account mechanisms coming from evolutionary biology are the Venus's-flytrap or automatic reflexes such as automatic removal of one's hand from the hot stove. The adaptive link account takes our *evaluative/generic* judgments as well as their "proto" versions to be such an adaptive link mechanism. For instance, a reflex mechanism is seen as a pairing between, for example, the circumstance of putting one's hand on the hot surface and the response of withdrawing one's hand. In the case of *evaluative* judgments, Street (2006, 127–128) gives an

example of reciprocal altruism in terms of pairing between the circumstance of being helped and the response of helping in return which has tended to promote one's survival and reproductive success. Similarly, in the case of *generic* judgments, one could argue that there is an adaptive link between the circumstance of the presence of sharks, tigers, mosquitoes and our response of a general avoidance of them that has tended to promote one's survival and reproductive success.[178]

Furthermore, Street (2006, 159–160) distinguishes between a mechanism of *effecting pairings* and a mechanism of *detecting* or *tracking* circumstances and responses. Even though the *detecting* or *tracking* mechanism can play the role of detecting the circumstance in which one is helped, e.g. detecting the presence of a hot surface, these are still two separate mechanisms. Street rightly points out that: "Our capacity for (non-evaluative) factual judgement does the job of tracking circumstances (tracking, among innumerable other things, which individuals have helped us), whereas our capacity for evaluative judgement does the job of effecting pairings of (perceived) circumstance and response (getting us, among many other things, to respond to those who have helped us with help in return)" (Street 2006, 160). Similarly, one may argue that our capacity for (non-generic) factual judgment does the job of tracking circumstances (tracking, among innumerable other things, that e.g. sharks eat swimmers, tigers are striped, mosquitoes carry the West Nile virus), whereas our capacity for generic judgment does the job of effecting pairings of (perceived) circumstance and response (getting us, among many other things, to avoid sharks, tigers, mosquitoes).

However, Street considers a potential objection based on the claim that there is a difference[179] between the mechanism of a reflex response and the mechanism of an evaluative judgment:

> The former is a brute, hardwired physical mechanism, while the latter is a conscious mental state, subject to reflection and possible revision in light of that reflection. But this does not change the fact that there is a deep analogy between their functional roles. From an evolutionary point of view, each may be seen as having the same practical point: to get the organism to respond to its circumstances in a way that is adaptive. (Street 2006, 128)

Street (2006, 128) suggest a further response to the above objection by arguing that in the case of a reflex mechanism a particular hard-wiring of the nervous system is responsible, while in the case of an evaluative judgment (or a more primitive evaluative experience including those that some other animals are likely to have) this happens "by having the organism experience a particular response as called for, or as demanded by, the circumstance in question. In the latter case, the link between

---

[178] A point that can be made about newly introduced evaluative and generic judgments is that, in some cases, it will take time before the adaptive link can be properly forged. In some other, such link may not form at all.

[179] It is worth noting, however, that neither Street (see 2006, 160) nor I are trying to offer a full explanation of why we evolved the "normative capacity" to make evaluative judgments or why we evolved the "generic capacity" to make generic judgments that ended up serving us for making links between circumstances and response, instead of getting this links forged by brute reflex mechanism. The tentative answer I adopt from Street is that evaluative and generic capacity are more flexible and plastic in comparison to the brute reflex mechanism. However, the answer to this question lies outside of the scope of the Darwinian dilemma.

circumstance and response is forged by our taking of the one thing to be a reason counting in favor of the other—that is, by the experience of normativity or value" (Street 2006, 128).

A similar response can be adopted in the case of an analogous objection to the adaptive link mechanism for generic judgments. In response, one could argue that there is a deep functional analogy between hardwired physical mechanism such as, for instance, a Venus's-flytrap, and a conscious mental state (that can be a subject to reflection and possible revision in light of that reflection) such as generic judgment. In particular, they can both serve the practical purpose of adaptive response of an organism to certain circumstances. Furthermore, in the case of generic judgments (or its proto-versions) organism experience a certain response depending on the particular circumstance, and the link between circumstance and response is forged by the experience of genericity.

5.5.2.3 Two Competing Scientific Accounts

As argued above, both the tracking relation account and the adaptive link account are evolutionary accounts of why we tend to make some *evaluative/generic* judgments rather than some others. More importantly, both of them rest on scientific explanations. As such, they are two competing scientific accounts that can be compared on scientific merits. Thus, in order to close a Darwinian dilemma against realist theories of *value/genericity* it remains to show that the adaptive link account for *value/genericity* offers a better scientific explanation than the tracking relation account for *value/genericity*, i.e. the one that is presuposing the independent *evaluative/generic* truths, respectively.

Before comparing the two scientific accounts, here is a quick summary of the crucial differences between the tracking and the adaptive link account Street offers for value which can also be adopted for genericity:

> According to the tracking account, however, making such evaluative judgements contributed to reproductive success because they are true, and it proved advantageous to grasp evaluative truths. According to the adaptive link account, on the other hand, making such judgements contributed to reproductive success not because they were true or false, but rather because they got our ancestors to respond to their circumstances with behavior that itself promoted reproductive success in fairly obvious ways… (Street 2006, 129)

In the rest of this section, assuming a direct analogy, I adopt and apply Street's arguments for why the adaptive link account provides a better scientific explanation of the same phenomenon than the tracking relation account, namely that the adaptive link account is superior by the usual criteria of scientific adequacy: (i) *parsimony*, (ii) *clarity*, and (iii) *shedding more light on the explanandum in question*.

Parsimony

The adaptive link account explains the existence of *evaluative/generic* judgments more

parsimoniously, without presuposing an extra ingredient, namely the *evaluative/generics* truths, in order to explain the phenomenon.

In comparison to the adaptive link account, the tracking relation account presupposes the additional ontological bit, namely it invokes the existence of the independent *evaluative/generic* truths. According to the tracking relation account, the independent truths are part of a scientific explanation, namely the tracking relation account presupposes that tracking such truths while making certain *evaluative/generic*s judgments rather than others is what benefits our survival and reproduction. On the other hand, the adaptive link account eschews positing the existence of truths entirely while explaining the benefits of making certain *evaluative/generic* judgments than others directly, by claiming that such judgments made creatures act in ways there were beneficial for their survival and reproduction.

Clarity

The adaptive link account is much *clearer* than the tracking relation account because it does not call for further explanations. The idea behind the tracking account is that because certain *evaluative/generic* judgments are true, grasping those truths via making certain judgments rather than some others promoted our reproductive success and survival.

However, the realist owes us a further answer *why* this is so beyond the one that it is so just because those judgments are true (see Street 2006, 129–130). The realist could offer a simple answer to this question by claiming that, as Street suggests, "on the face of it, it might seem that of course it promotes reproductive success to grasp any kind of truth over any kind of falsehood. Surely, one might think, an organism who is aware of the truth in a given area, whether evaluative or otherwise, will do better than one who isn't" (Street 2006, 130). However, Street (2006) quickly dismisses this line of thought by showing that it would actually be disadvantageous from an evolutionary point of view to posit such independent truths. She considers two kinds of truths. On the one hand, the truths such as the truths about a creature's manifest surroundings, e.g. a fire raining in front of it, predator rushing towards it, for which we can see how grasping them will promote their reproductive success and survival. On the other hand, the truths that will confer either no advantage or a disadvantage to grasp them, e.g. truths about the presence or absence of electromagnetic wavelengths of the lowest frequencies. One may argue that, according to the realist, *evaluative/generic* truths are more like the later kind of truths. If so, the realist again needs to explain in what way grasping evaluative/generic truths that are more like the later ones would promote our reproductive success and survival. However, it turns out that the standard evolutionary answer is not available for the independent *evaluative/generic* truths. As Street puts it:

> To say that these truths could kill you or maim you, like a predator or fire, would be one kind of answer, since it makes it clear how recognizing them could be advantageous. But such an answer is clearly not available in the case of the independent (…) truths posited by the non-naturalist realists. In the absence of further clarification, then, the non-naturalist's version of the tracking account is not only less parsimonious but also quite obscure. (Street 2006, 131)

Moreover, apart from lacking a clear explanation about why we would develop such independent truths, Street (see 2006, 130) further claims that, from the evolutionary perspective, developing an ability to grasp such independent truths would be spending our energy and resources.[180]

Furthermore, one can consider whether the *value/genericity naturalist* would run into similar problems (see Section 5.4.3.3). At first glance, the value/genericity naturalist's view seems more promising since *value*/genericity naturalist construes *evaluative*/*generic* facts as natural facts with causal powers, and grasping such truths could have an impact on our reproductive success and survival. However, one could raise a similar worry for the *value*/genericity *naturalist*, namely one could ask what kind of natural facts are we talking about and why grasping them promotes our reproductive success and survival? The value naturalist does not have a good answer either. Following Street, one could argue that the explanation about exactly "what natural fact or facts does the evaluative fact that one should care for one's offspring reduce to, or irreducibly supervene upon, and why would perceiving the natural fact or facts in question have promoted our ancestors' reproductive success" brings in complexities that the adaptive link account avoids by saying "that ancestors who judged that they should care for their offspring met with greater reproductive success simply because they tended to care for their offspring—and so left more them" (Street 2006, 131–132).

A similar point can be adopted for generic truths. The naturalist about genericity needs to offer an explanation for exactly what kind of natural fact or facts does the generic fact (that e.g. tigers have stripes) reduce to, or irreducibly supervenes upon, and why perceiving the natural fact in question would have promoted our ancestors' reproductive success. The answer brings in complexities that, on the other hand, the adaptive link account avoids by saying that ancestors who judged that tigers have stripes met with greater reproductive success simply because that helped them to recognize tigers and avoid them which, in result, promoted their survival and reproductive success.

Explaining the Explanandum

The explanandum at hand is that humans have widespread tendencies to make certain *evaluative*/*generic* judgments rather than some others. The adaptive link account is more illuminative with respect to why we have these tendencies than the tracking relation account. Here is Street's brief comparison between the answers that the two accounts in question can offer:

[180] Notice, however, that this still does not mean that Street can argue that this is an explanation that such independent truths do not exist.

> The adaptive link account has very good answers to such questions, of the general form that ancestors who made evaluative judgements of these kinds, and who as a result tended to respond to their circumstances in the ways demanded by these judgements, did better in terms of reproductive success than their counterparts. (…) Now compare the tracking account's explanation. It tries to answer these same questions by saying that these judgements are *true*: that survival *is* valuable, that we *do* have special obligations to care for our children, that the killing of human beings *is* more serious than the killing of plants or other animals. Such answers do not shed much light. (Street 2006, 132)

Similarly, one could argue that the adaptive link account for genericity offers an answer why we have tendencies to make generic judgments by claiming that the ancestors who made generic judgments, and who as a result tended to respond to their circumstances in the ways demanded by these judgments, did better in terms of reproductive success than their counterparts. On the other hand, one could argue that the tracking relation account for genericity could explain these tendencies by merely saying that these judgments are *true*: e.g. that tigers have stripes is true, that mosquitos carry the West Nile virus is true. However, the latter explanation does not shed much light.

In the rest of this section, I consider three questions (see Street 2006, 133–135) with respect to explaining the explanandum that the tracking relation account for evaluative and the tracking relation account for generic truths fail to answer.

Question One: How does the tracking relation account explain the remarkable coincidence that so many of the truths it posits turn out to be exactly the same judgments that forge adaptive links between circumstance and response—the very same judgments we would expect to see if our judgments had been selected on those grounds alone, regardless of their truth?[181]

The answers that the tracking relation accounts, both the one for evaluative and the one for generic truths, can offer will be confronted with the parsimony and clarity problems that I have discussed above.

Question Two: What does the tracking relation account have to say about our observed predispositions to make other *evaluative/generic* judgments which (we may decide on reflection) are *not* true?[182]

Street (see 2006, 133) gives an example of our deep tendency to treat those from the "in-group" better than those from the "out-group." She contrasts the two accounts and argues that the adaptive link account provides a better explanation for this deep tendency by claiming that this tendency tends to promote our reproductive success in those who possessed it since they tend to help those that are more genetically related to them, or those that are more able or likely to reciprocate the help. On the other

[181] I adopt this question from Street (2006, 132).

[182] I adopt this question from Street (2006, 133).

hand, the tracking account's explanation that this deep tendency developed because "it is true that the fact that someone is in an out-group is a reason to accord him or her lesser treatment than those in the in-group" is not plausible, especially given that we increasingly believe that this is not true (see Street 2006, 133).

I believe this point is especially valid in the case of generic judgments, and, in particular, in the case of a subclass of generic judgments that involve social group terms such as "Muslims are terrorists," "Blacks are violent," "Women are submissive," "Girls are bad at math," etc. The adaptive link account for genericity can offer a promising explanation of these generic tendencies, namely that having these generic tendencies tended to promote our reproductive success while also being compatible with why we tend to reject these generics upon *rational reflection*.

Depending on the content of a particular generic judgment we could offer different evolutionary explanations for having those generic tendencies. For instance, the evolutionary explanation for generic judgments such as "Muslims are terrorists," "Blacks are violent," could, at least in part, fall back onto our tendencies to treat those from the "out-group" lesser than those from the "in-group," whereas our tendencies for generic judgments such as "Women are submissive," "Girls are bad at math," could be, at least in part, explained by using the evolutionary explanation of how the roles that females of certain species occupied in the given circumstances lead to promotion of the reproductive success of that species.

On the other hand, the tracking relation account for genericity could only offer an implausible answer that this generic predisposition developed because it is *generically true* that Muslims are terrorists, Blacks are violent, women are submissive, girls are bad at math, etc. In fact, both the theorists as well as the speakers are increasingly coming to think that this is not *generically true*. The tracking relation account for genericity is, thus, left with nothing in the way of an explanation as to why we observe such deep tendencies to make the contrary judgment. In a similar way as in the case of some of our evaluative judgments, many of us are increasingly coming to think that these generic judgments are not correct (see Chapter 6).[183]

Finally, Street considers another potential objection according to which the advocate of the tracking account could argue that they have a better explanation of our *corrected* evaluative judgments, like a corrected judgment that being a member of "out-group" is not a reason to treat that person less than the person that is a member of "in-group." However, Street argues that corrected evaluative judgments are perfectly compatible with the adaptive link account, given that one can come to reject some of her basic evaluative tendencies on the basis of other evaluative judgments she holds (see Street 2006, 161).

---

[183] Philosophers have recently offered different accounts of why such judgments may not be admissible or correct, e.g. they might be semantically false, they are falsely essentializing certain kinds, they are fostering socially, politically or morally pernicious effects (see Chapter 6).

One could offer a similar answer to the potential objection that a tracking relation account for genericity gives a better explanation of our *corrected generic judgments*, e.g. our corrected judgment that Blacks are not violent or that women are not submissive. In particular, one could, similarly, claim that rejecting some of our basic generic tendencies on the basis of other generic (and, perhaps, other, non-generic) judgments we hold is compatible with the adaptive link account.

Question Three: How come that there are all those *evaluative/generic* judgments that human beings could make but do not?[184]

The universe of logically possible, both evaluative as well as generic, judgments is huge, yet we do not have tendencies to make certain evaluative/generic judgments such as, for example, the evaluative judgment "that infanticide is laudable, to the judgement that plants are more valuable than human beings, to the judgement that the fact that something is purple is a reason to scream at it" (see Street 2006, 133) or the generic judgment that humans have green bodies, sharks sing beautiful songs, ravens are bigger than toasters.[185] As Street notices:

> Here again the adaptive link account has something potentially informative to point out, namely, that such judgements—or evaluative tendencies in these general sorts of directions—forge links between circumstance and response that would have been useless or quite maladaptive as judged in terms of reproductive success. The tracking account has nothing comparably informative to say. It can just stand by and insist that such judgements are false—reaffirming our convictions but adding nothing to our understanding of why we have them. (Street 2006, 133)

One can make a similar claim for our generic tendencies, namely the advocates of the adaptive link account could say that had we had generic tendencies and judgments of the above mentioned sort, forging links between a circumstance and a response would have been useless or quite maladaptive as judged in terms of reproductive success. The tracking relation account for genericity does not have much to say apart for claiming that such generic judgments are false, without having an explanation why we have them.

To conclude, let me generalize Street's (see 2006, 134–135) points about value, in order to close a Darwinian dilemma for realist theories of *genericity*. The tracking relation account is the most natural account the realist can give in order to explain the relation between *generic* judgments (under the assumption that they have been influenced by evolutionary forces) and the independent *generic* truths that the realist posits. For, given the nature of realism itself, all the alternative accounts the realist may offer are, once again, forced to give the tracking relation account since that is the only way for the realist to assert the above mentioned relation without stepping back into the first horn of the dilemma which

---

[184] I adopt this question from Street (2006, 133).

[185] "Raven's are bigger than toasters" is a generic used by Nickel (2016, 7).

implausibly denies such relation. As a result, both denying the relation as well as accepting the relation seem untenable.

5.6 Conclusion

In this chapter, I have argued for a Darwinian dilemma for realist theories of *genericity* based on a direct analogy with Sharon Street's (2006) Darwinian dilemma for realist theories of *value*. In particular, I have shown that there is a parallel we can draw between the relation between our evaluative judgments and generic judgments on the one hand, and realism about value and realism about genericity on the other. Given this parallel, I have further shown that a Darwinian dilemma that applies to the realist theories of value applies to the realist theories of genericity as well.

This chapter, thus, questions our reason to believe in the existence of the independent generic truths or facts. It works under the assumption that speakers' generic judgments are responsive to something that is wholly independent of the objective generic facts/truths the realist posits. In particular, I have argued that realist theories of genericity prove unable to accommodate the fact that Darwinian forces have deeply influenced the content of human genericity. In result, my argument aims to remove the evidential force of generic judgments, which have been thought to support generic truths.

More broadly, this chapter can be seen as a pushback against the realist theories of genericity, initially introduced in order to solve the problem of giving a semantics for generics. Instead of expecting the unity to arise from the theory of genericity, this chapter utilizes a plausible hypothesis that our generic judgments have been influenced by evolutionary forces in order to offer an explanation of why the set of generic judgments we affirm seems to be so unrelated in their content.

Moreover, I believe that a Darwinian dilemma for realist theories of genericity offers some of the motivational force for antirealism about genericity (see Fuš manuscript a), as well as for engineering generics at the level of judgments (see Chapter 6). The adaptive link account can be used to explain an Alternative Assumption of Unity,[186] namely the one offered by the antirealist about genericity which does not need to presuppose the existence of genericity in the realist sense.

Finally, this chapter offers another way to take evolutionary psychological research seriously. Instead of packing the psychological research into the theory of genericity (see Leslie 2007, 2008, 2017), this chapter shows that we can utilize the psychological research against the realist theories of genericity.

[186] Recall (Section 5.3.3), the realist about genericity posits: "Assumption of Unity: There is a unified phenomenon of genericity that generic sentences, in general, instantiate" (Sterken 2015, 2).

CHAPTER 6

THE SPEECH ACT APPROACH TO ENGINEERING GENERIC JUDGMENTS

Abstract

In this chapter, I address the following question: Can the pernicious effects of generic statements such as "Blacks are violent," "Women are submissive," or "Muslims are terrorists" be ameliorated by applying the method of engineering in philosophy? My contribution to the debate about the so-called pernicious generics is twofold: (i) I identify and critically examine current approaches to pernicious generics (Haslanger 2011; Leslie 2017; Saul 2017; Ritchie 2019) as domain-specific instances of philosophical engineering of generics; (ii) I develop an alternative domain-specific account of philosophical engineering of pernicious generics. In particular, I argue that the pernicious effects of generics stem most fundamentally from our judgments about generic propositions. As one of the plausible strategies to achieve engineering of generic judgments, I suggest the Speech Act Approach, an approach to changing doxastic propositional attitudes towards generics by utilizing speech acts. I show that the Speech Act Approach has three key benefits. Firstly, it does not depend on any particular view about the semantics, epistemology or metaphysics of generics. Secondly, it is more feasible to implement than the proposed alternatives. Thirdly, it provides a better framework for utilizing generics in order to achieve beneficial or neutralizing effects. Finally, I sidestep an ethical/epistemic dilemma for the Speech Act Approach to engineering generic judgments.

6.1 Introduction

Generics are ubiquitous in natural languages. They are considered to be an important source of information both for adults as well as for preschoolers since they allow for a quick and easy way to convey generalizations without the need to account for exceptions.[187] In many cases, generics are perceived as *beneficial* generalizations. For example, consider generics such as "Tigers have stripes," "Mosquitos carry the West Nile virus," "Ducks lay eggs," "The Dodo is extinct," etc. However, certain generics have been perceived as *pernicious* generalizations. For example, consider generics such as "Blacks are violent," "Muslims are terrorists," "Latinos are lazy," "Women are submissive," "Girls like pink," "A woman puts family before career," etc. The latter generics raise pressing issues because they can, arguably, foster morally, socially or politically *pernicious* effects by, for example, reinforcing harmful or false stereotypes. In the recent years, there has been a burgeoning interest in addressing the pernicious effects of generics. In particular, philosophers have focused on explaining the *sources* and offering *remedies* for their perniciousness. The general aim of this chapter is to contribute to philosophical discussions about the so-called *pernicious generics* and to elucidate their connections to *philosophical engineering*.

My contribution to the debate about the so-called *pernicious generics* is twofold: (i) I identify and critically examine current approaches to pernicious generics (Haslanger 2011; Leslie 2017; Saul 2017; Ritchie 2019) as domain-specific instances of philosophical engineering of generics;[188] (ii) I develop an alternative domain-specific account of philosophical engineering of pernicious generics. In particular, I argue that the pernicious effects of generics stem most fundamentally from our judgments about generic propositions, i.e. generic judgments. As one of the plausible strategies to achieve engineering of generic judgments, I suggest the Speech Act Approach, an approach to changing doxastic propositional attitudes towards generics by utilizing speech acts. I show that the Speech Act Approach has three key benefits. Firstly, it does not depend on any particular view about the semantics, epistemology or metaphysics of generics. Secondly, it is more feasible to implement than the proposed alternatives. Thirdly, it provides a better framework for utilizing generics in order to achieve beneficial or

---

[187] "By 30 months, children understand that generics tolerate exceptions (Gelman and Raman 2003), and several studies have found that children's judgments of generics are similar to adults' judgments throughout the preschool years (e.g., Brandone, Cimpian, Leslie and Gelman 2012; Brandone, Gelman, Hedglen 2015; Hollander, Gelman and Star 2002). Interestingly, preschool children, like adults, understand that, e.g., "birds lay eggs" can be true yet "birds are girls" false, even though only "girl" birds lay eggs. Thus even preschoolers understand at least some of the "troublesome" generics—the generics that cause difficulties for the standard semantic analyses of generics—in the same way that adults do." (Leslie and Lerner 2016)

[188] Throughout this chapter, I reconstruct and evaluate methodological frameworks behind the main current approaches within the debate about pernicious generics. Despite the lack of authors' explicit mention of the method behind the current approaches to pernicious generics, my diagnosis is that these accounts are attempts at engaging in domain-specific instances of philosophical engineering (for support of the idea that conceptual engineering is not a "luminous condition," i.e. one can engage in it without knowing that one is engaging in it, see Cappelen (2018, 73)). In particular, I show that the participants in the current debate about pernicious generics submit to two core theses characteristic of the method of philosophical engineering*: (a) Evaluation Thesis:* Generics can have (morally, politically, socially) pernicious effects, and (b) *Amelioration Thesis*: We ought to find remedies for these effects.

neutralizing effects. Finally, I sidestep an ethical/epistemic dilemma for the Speech Act Approach to engineering generic judgments.

Since my main interest in the debate about pernicious generics comes from the philosophical engineering angle, a few quick caveats are in order. The discussions of the underlying literature on semantics and metaphysics of generics are kept down to the minimum (I covered them in more detail in Chapter 5). More importantly, this chapter should be read as a case study of Part I of this dissertation. Its structure reflects the potential of applying the Five Stage Recursive Model of the method of philosophical engineering within the debate of pernicious generics. Along the way, I illuminate a trend in the recent literature on pernicious generics that, I believe, can be seen as falling under the method of philosophical engineering. However, the critical material of the particular views in the current debate is mainly in the service of creating space for an alternative approach that I introduce and, thus, calls for a more elaborate and independent discussion.

Here is how I proceed. Throughout Sections 6.2–6.6, I apply the *five stage recursive model of philosophical engineering,* that I developed in Part I of this dissertation (see Chapter 2, Section 2.3), to generics. In particular, in Section 6.2, I *identify* the main target of engineering generics; in Section 6.3, I critically examine current and offer my own take on *evaluation* of generics; in Section 6.4, I critically examine current and offer an alternative *strategy* for engineering generics; in Section 6.5, I offer some general considerations for *implementation* of my strategy; in Section 6.6, I briefly stress the main roles of *re-evaluation*. In Section 6.7, I outline the three key *benefits* of my proposal, and in Section 6.8, I diffuse an *ethical/epistemic dilemma* surrounding my approach.

6.2 Identification

There has been a general agreement among the theorists that certain generics can lead to pernicious effects (see Haslanger 2011; Langton, Haslanger, and Anderson 2012; Leslie and Tworek 2012; Wodak, Leslie, and Rhodes 2015; Rhodes, Leslie 2017; Saul 2017; Wodak and Leslie 2017; Ritchie 2019). Recent debate on pernicious generics focuses largely on a sub-class of *characterizing generics* (see Chapter 5, Section 5.2), the so-called *striking property generics* (see Leslie 2014, 2017). The main characteristic of *striking property generics* is that speakers tend to *accept* certain generics *as true* at a low prevalence[189] when the predicated property in question is *dangerous* or when they have strong interest in *avoiding* the predicated property.

Psychological research has shed some important light on this phenomenon. Leslie (2017, 396), relying on her previous work on the cognitive structure and semantics of generics (see Leslie 2007,

[189] I.e. despite very few members of the kind in question possessing the predicated property. For example, speakers *assert* "Mosquitoes carry West nile Virus" (while currently less than 1% of actual mosquitoes do), "Sharks attack swimmers," "Blacks are rapists,""Muslims are terrorists," yet *refuse* "Books are paperbacks" (while currently around 70% of actual books are) and "Office chairs explode."

2008), argues that: "The disposition to generalize strikingly negative information on the basis of even a single event, thus, appears to be a pervasive aspect of our thinking" (Leslie 2017, 396). As a result of this disposition, such generalizations turned out not to be "psychologically inert" but also play an important role when it comes to our *inferences concerning property possession* and *guide our judgments about members of a kind* (see Leslie 2017, 398; see also Section 6.3.1).

Striking property generics may serve *beneficial* purposes. For example, consider the potential benefits for our survival by accepting claims such as "Mosquitoes carry the West Nile virus," "Sharks attack swimmers," "Pit bulls maul children," "Deer ticks carry Lyme Disease," "Tigers eat people," etc. However, striking property generics, especially those about *social groups* such as "Blacks are violent," "Muslims are terrorists," "Latinos are lazy," "Women are submissive," etc., are perceived as particularly *pernicious*. It has been argued that striking property generics about social groups can contribute to (e.g. morally, socially or politically) *pernicious effects* such as the perpetuation of (possibly false) racist, pejorative, religious or sexist essentialist beliefs which can lead to prejudiced attitudes and behavior towards the respective social groups. It has also been argued that this can further result in e.g. systematic patterns of violence, discrimination, (negative) stereotyping and stereotype threat, hate speech, implicit biases, structural oppression, social injustice regarding social groups that are stigmatized as dangerous, etc. (see Haslanger 2011; Gendler 2011; Langton, Haslanger, and Anderson 2012; Leslie 2017; Saul 2017; Ritchie 2019).

Even though *the main target of the current projects* about pernicious generics are the above mentioned pernicious effects connected to the *striking property generics about social groups, three caveats* about the scope of the target of this chapter are in order.

First. Generics that are *not striking property generics* (e.g. "A woman stays home and raises a family," "A woman puts family before career," "Girls like pink," "Women belong to the kitchen," "Latinos are temperamental," "Women are nurturing," "Asians are smart," "Blacks are good at basketball") and generics that are *not about social groups* (e.g. "Sharks eat swimmers" or "Pit bulls maul children") could, arguably, lead to similar pernicious effects, though potentially with a lower rate of acceptance for the reasons mentioned above.

Second. This chapter operates under the assumption that given the right conditions many (if not all) generics could be perceived as pernicious since the *valence* (i.e. perniciousness, benefits, or neutralness) of the effects is largely connected to *non-generic facts* (see Chapter 2, Section 2.3.2.2). It is, however, beyond the scope of this chapter to investigate all of the *non-generic facts* that play a role in the perniciousness of generics and their interconnection with generic facts. Instead, here is a brief illustration. Whether a certain generic statement will turn out pernicious is not restricted only to a particular generic but can instead depend on factors such as underlying social, moral and political

norms and practices. For instance, if one regarded animal rights very highly and wanted to protect their lives as well as their dignity, one could argue that our use of *striking property generics about natural kinds* such as "Sharks eat bathers," "Pit bulls maul children," or "Ticks carry Lyme virus" could lead to bad consequences for these animals. It is possible to imagine that some of the bad consequences of the above mentioned generics could lead to discrimination and fostering prejudice as well as stereotypes about the above mentioned animals, in addition to an increased endangerment of their lives (since those who see them as a threat might be more prone to kill them).[190]

Third. In this chapter, I am primarily concerned with the above mentioned *pernicious effects that stem from generic propositions*. However, it is important to notice that generics are neither *necessary* (i.e. the above mentioned pernicious effects such as e.g. stereotyping or hate speech may be achieved without the use of generics) nor *sufficient* (i.e. mere use of a generic proposition will not in all cases bring about the above mentioned pernicious effects) for bringing about the above mentioned pernicious effects. As a corollary, ameliorating pernicious effects connected to generic propositions is only one way to fight the above mentioned pernicious effects. Other kinds of engineering projects (such as e.g. social, moral or political engineering) may fight the same kind of pernicious effects more broadly, and often in synergy with different forms of philosophical engineering (for *engineering holism* and *division of engineering labour* more generally see Chapter 2, Section 2.3.3; and in connection to generics in particular see Section 6.4.5, and Section 6.5.4).

6.3 Evaluation

In the *identification* stage, as the main target of the present engineering project, I identified the pernicious effects of generics. In the evaluation stage, I am further examining the main source of these effects, i.e. how is it that generics have these effects. In this section, after discussing some of the current proposals (Sections 6.3.1–6.3.4), I offer my own take on the evaluation of what brings about the effects of generics (Section 6.3.5).

6.3.1 Leslie

Leslie (2017, 404) formulates the *truth-conditions* of striking property generics as follows: "A generic statement in which a striking property is predicated is, I claim, true if and only if some members of the kind in question possess the relevant property, and the others are typically disposed to possess it" (Leslie 2017, 404). More importantly, Leslie argues that certain striking property generics about social groups (interestingly enough, the pernicious ones) are *false*. According to her view, striking property generics such as "Muslims are terrorists" and "Blacks are violent" are *false* because "given half a

[190] See "PETA just compared using animal idioms to racism and homophobia and the backlash is real" (Independent 2019).

chance" Muslims are not disposed to be terrorists and Blacks are not disposed to be violent (see Leslie 2017).

Leslie offers truth-conditions of striking property generics in a way that goes against the speakers' intuitions[191] about the truth value of the above mentioned generics (see also Section 6.7.1). In order to explain why speakers assert these generics contrary to their truth value, Leslie (2017) utilizes data from cognitive and evolutionary psychology (see Hollander, Gelman and Star 2002; Leslie 2008, 2017; Rhodes, Leslie, and Tworek 2012). She develops a psychological approach to generics which operates under the assumption that *our judgments* of generic propositions may be sensitive to a range of factors connected to the content of a particular generic, such as a *predication of a striking property in the presence of a kind term*. For instance, according to this approach, when a *kind term* and *predication of a striking property* occur together in a form of a generic proposition, the *assertion rate* of that particular generic is higher. Leslie claims that this is due to the phenomenon called '*psychological essentializing*' (as opposed to *metaphysical essentializing*) (see Leslie 2013, 2014, 2017). Psychological essentializing is thought to occur both in the presence of kind terms such as *tiger* that belongs to a *biological domain* as well in the presence of kind terms such as *gender* and *sex* from a *social domain*.[192] For the social realm to be essentialized in a psychological sense it means that "people believe that members of these categories share a fundamental nature that grounds a range of common properties" (Leslie 2014, 4).

Leslie uses psychological essentializing as an explanation for why certain generics are *accepted* contrary to the prevalence of the property among members of the kind. She draws on psychological research that suggests that from an early age speakers *take* certain kinds *as* having fundamental natures that all the members share (see Gelman 2003; Leslie 2014). As Leslie (2014) puts it, speakers believe that:

> … certain kinds 'carve essence at its joints'—that is, certain kinds (e.g. animal kinds) pick up on genuine differences and similarities in the essences of individuals. Thus, speakers believe that the kind tiger picks out individuals with highly similar essences, and these shared essences explain why tigers share so many outwardly observable properties, such as having stripes and tails and being ferocious. (see Leslie 2014, 4)[193]

Leslie (2017) argues that in the case of striking property generics fewer instances of the property can be possessed by members of the kind in order for speakers to judge the generic to be true (when compared with non-striking property generics). It takes far fewer instances of, for example, *murder* for one to be considered a murderer than it does instances of *anxiety* to be considered a worrier. Most importantly, Leslie (2017) claims that psychologically essentialized social groups tend to be more likely

---

[191] For more about Leslie utilizing a debunking argument to argue for their truth value see Fuš manuscript a,b.

[192] See Rothbart and Taylor 1992; Hirschfeld 1996; Gelman 2003; Prentice and Miller 2007; Rhodes and Gelman 2009; Meyer et al. 2013.

[193] For more details see Gelman 2003; Leslie 2013.

to be stereotyped and prejudiced from both children and adults, and that striking property generics about social groups may illuminate some social stereotypes (e.g. "Black men are rapists") better than others (e.g. "Black men are athletic).[194]

In a nutshell, Leslie (2017) argues that speakers who believe that a generic such as "Muslims are terrorists" is true make a *mistake* when they take Muslims as a *psychologically essentialized* kind, namely when they think that the social kind Muslim is a good predictor of possessing the striking property attributed by the generic, i.e. the property of 'being a terrorist.' In result, this leads to the striking property generic "Muslims are terrorist" to be *accepted* as true by the speaker. Leslie's approach, as Cappelen (2018) points out, is rather unique since "she presents empirical evidence that the use of certain linguistic constructions lead those using them to make cognitive mistakes. The expressions in question are the so-called generics, and the mistake in question is that of essentializing social kinds" (Cappelen 2018, 18). In other words, Leslie (2017) believes that in the case of striking property generics about social groups speakers often make a *cognitive mistake of psychological essentializing* resulting in *false predication* of a certain property (e.g. being a terrorist) to particular members of a kind (e.g. Muslims) that do not possess the predicated property in question (and, according to her, they also do not have a disposition to possess the predicated property and it is not in their essence to possess it).

Cappelen (2018, 33–34) takes this cognitive mistake of generics to fall under what he calls '*objectionable effects of the semantic value.*'[195] However, in the present context, 'the semantic value,' according to Cappelen (2018), arguably refers to the presence of a predication of a striking property occurring in certain social kind generics. This can further contribute to a cognitive effect of psychological essentializing which can contribute towards higher rates of assertion that can lead to pernicious effects. Thus, what Cappelen categorized as the *objectionable effects of the semantic value* should, more precisely, be called the *objectionable effects of assertion governed by psychological essentializing* (whereas psychological essentializing is the effect of the semantic value).

However, it is important to notice that even if Leslie (2017) is correct to connect the phenomenon of

---

[194] As Leslie (2014, 212–213) points out: "[psychologically] essentialized social categories are more likely to be the targets of sustained and virulent prejudiced attitudes (see e.g. Haslam et al., 2000, 2002), and Keller (2005) offers experimental evidence that suggests there may indeed be a causal link between essentialist beliefs and prejudice. Relatedly, essentialized groups tend to be more susceptible to stereotyping (e.g., Bastian and Haslam, 2006; Prentice and Miller, 2006, 2007; Williams and Eberhardt, 2008; Yzerbyt et al., 2001)" (Leslie 2014, 212–213). However, it is important to notice that psychological essentializing does not need to lead only to pernicious effects of generics. In fact, as I discuss in Section 6.7, it can be utilized to bring about beneficial effects (see also Saul 2017, Ritchie 2019).

[195] As part of his taxonomy of conceptual deficiencies, Cappelen (2018) considers a deficiency according to which *semantic value* itself is not defective, i.e. a particular expression "has a [non-deficient] semantic value, but for that particular expression to have that semantic value has bad effects" (Cappelen 2018, 34). He dubs this deficiency 'objectionable effects of the semantic value' (Cappelen 2018, 34), and distinguishes between different objectionable effects of the semantic value: *morally, politically, or socially objectionable effects, cognitive effects*, and *effects on theorizing* (Cappelen 2018, 33–34) (see also Chapter 2, Section 2.3.2.2).

psychological essentialising to the higher rates of *assertion*[196] of certain false generics and, possibly, false predication, the truth value of a generic proposition does not *per se* contribute to psychological essentializing. Furthermore, psychological essentializing, according to Leslie, is the reason why speakers *violate the truth assertion norm*, i.e. assert false generics. However, *conforming to the truth assertion norm* and *psychological essentializing* are neither *sufficient* nor *necessary* for the effects of generics. *Conforming to the truth assertion norm* is not *sufficient* because there are cases where conforming to the truth assertion norm does not lead to pernicious effects of generics. Conforming to the truth assertion norm is also not *necessary* because pernicious effects of generics can occur without c*onforming to the truth assertion norm*. Similarly, *psychological essentializing* is not *sufficient* because there are cases of psychological essentializing which do not have pernicious effects. Psychological essentializing is also not *necessary* because, the pernicious effects of generics can occur without psychological essentializing.

Relatedly, if Leslie is wrong[197] about the truth-conditions of striking property generics and it turns out that the above generics such as "Blacks are violent," "Women are submissive," or "Muslims are terrorists" are, in fact, *true*, conforming to the truth assertion norm means that the speakers should *assert* generics such as "Blacks are violent," which would lead to pernicious effects, the effects Leslie wanted to avoid in the first place. On the other hand, even if Leslie is right and the above generics such as "Blacks are violent," "Women are submissive" or "Muslims are terrorists" are *false*, conforming to the truth assertion norm is *supererogatory* or *irrelevant* for *brining about* or *getting rid* of moral, social and political effects, given that it is not the truth but rather the assertion that influences our generic judgments (see Section 6.3.5).

At the risk of spoiling the suspense, here is a snapshot preview of my approach that takes this overlooked, nevertheless, important lesson about the relation between *an assertion, psychological essentializing* and *the effects of generics* seriously. In Section 6.3.5, I will argue that even though the *effects of the assertion governed by psychological essentializing* can lead to *pernicious generic judgments*, psychological essenitalizing is not *necessary* for the assertion of a certain generic to occur but can instead *influence the rate of assertion*. In other words, assertions that the speakers make in the presence of some generics is merely *facilitated* rather than *caused* by psychological essentializing. Instead, I argue that

---

[196] Speakers can *assert false* generics and *refuse true* generics. Semantics and pragmatics of generics can come apart (see Haslanger 2011). Furthermore, under the assumption that semantics of generics *tracks* metaphysical essentialism and that the pragmatics of generics *tracks/correlates with* psychological essentialism, there is a room for a potential mismatch, namely it is possible that psychological essentialism and metaphysical essentialism come apart.

[197] Saul (2017) and Sterken's (2015) raised worries about Leslie not offering plausible truth-conditions for striking property generics. Saul (2017) brings out a doubt about Leslie's consistency when it comes to the truth-conditions of striking property generics. In particular, Saul objects: "This is actually quite problematic for her when it comes to (1) [Mosquitoes carry the West Nile virus], which she wants to maintain is true. Mosquitoes are not disposed to carry West Nile virus 'given half a chance'—they need to be exposed to it" (Saul 2017, 4). However, even though it weakens Leslie's position, Saul's objection might not be entirely in place, because Leslie (2017, 404) admits that "perhaps only mosquitoes with a particular mutation are capable of carrying the virus. If these turn out to be the facts, then my account predicts that the above generics are in fact false and it is only the weaker claims [… such as] 'mosquitoes with a particular mutation carry the West Nile Virus' that are true" (Leslie 2017, 404).

speech acts, such as the speech act of *assertion,* can lead to a generic judgment resulting in pernicious effects. In Section 6.4.5, I motivate and develop an approach to achieve *amelioration of pernicious generic judgments* via utilizing *speech acts* connected to pernicious generics.

6.3.2 Haslanger

Contrary to Leslie's (2017) suggestion, Haslanger (2011) argues that the main source of perniciousness of generics is not connected to *false predication* but rather to *false implication*. Moreover, Haslanger (2011) wants to stay neutral on the truth-conditions of striking property generics. She claims that, when it comes to the truth-conditions of generics, speakers are not very good at thinking about them, and they often think that generics such as "Blacks are violent" or "Women are submissive" are *true*.

Haslanger (2011), however, does not want to exclude a possibility that (at least) some of the pernicious generics such as "Blacks are violent" or "Women are submissive" may be *true*. For instance, one could argue that she might be right to think that "Blacks are violent" could be true "because there might be a good explanation of black violence as a response to racist oppression, and this might supply a non-accidental connection that is sufficient to make (5) true" (Saul 2017, 4). Similarly, one could argue that Haslanger might be right to think that "Women are submissive" could be true in case "society punishes assertiveness in women to such an extent that they rarely, if ever, are other than perfectly submissive. It would be hard to argue that the generic "women are submissive" is then false, but there nonetheless would seem to be something damaging about asserting it" (Leslie 2014, 9).

Thus, no matter whether generics such as "Blacks are violent" or "Women are submissive" are true or false, Haslanger (2011) argues, they are *misleading* because even if striking property generics were semantically true, such generics carry *false implicatures about essence* that can lead to pernicious effects. As Leslie (2014) puts it:

> Haslanger proposes that the damage comes, at least in part, from our tendency to suppose that generic generalizations obtain because of inherent features of the kind in question. That is, even if "women are submissive" is made true by purely external sociological factors, it still may by default communicate that there is something in the nature of women that makes them submissive. That is, it may reinforce essentialist beliefs about women, and further communicate that this shared essence causes women to be submissive. (Leslie 2014, 9)

Haslanger, thus, sees the peril of generics such as "Blacks are violent" or "Women are submissive" in their *false conversational implicatures* which can easily become the background assumptions and can then be taken for granted. As such, Haslanger believes, they become hard to challenge and can hence contribute to pernicious effects.[198]

[198] Also, see Langton, Haslanger and Anderson (2012, 764).

I agree with Haslanger (2011) that *false predication* of generics is not the main candidate for the effects of generics (see Section 6.3.1). Moreover, I also agree that effects of generics are connected to pragmatic properties of generics rather than to their semantic properties (see Section 6.3.5). However, I don't believe that *false conversational implicature*, in particular, is responsible for effects connected to generic propositions. Instead, I share Saul's (2017) reservations about what kind and whether conversational implicatures are connected to the perniciousness of generics. In Section 6.3.5, I argue that a (false) implicature is not necessary for the effects of generics.

6.3.3 Saul

Saul (2017) takes a different approach to the evaluation of the pernicious generics than both Haslanger (2011) and Leslie (2017). In her paper titled 'Are Generics Especially Pernicious?', Saul's (2017) main claim is that *there is nothing special about generics.*

Against Haslanger (2011), Saul argues that "it seems unlikely that a claim about natures is conversationally implicated" (Saul 2017, 5). Instead, she argues that perniciousness is a wider phenomenon, namely that the main culprit is the *relevance implicature* which is, nevertheless, *also* present in cases which do not include only generics but all cases of mentions of racial, ethnic and other groups:

> Almost any time a race or ethnicity is mentioned, there will be an implicature that it's relevant to the subject matter under discussion. More generally, if specific groups of people are introduced into a discussion there will be an implicature that these specific groups are relevant to the discussion. Introducing racial, ethnic and others groups into discussions in this way undoubtedly makes a contribution to the transmission of pernicious ideologies. (Saul 2017, 6)

Against Leslie (2014, 2017), Saul argues that the pernicious statements such as "Many Asian men are abusive to women" and "My friend has a friend who's Asian and he's abusive to his wife" that are *not generic* can also carry the "same sort of relevance implicature" (Saul 2017, 6). As she puts it:

> I found these utterances very disturbing: it's undoubtedly true that many Asian men are abusive to women, as are many men from any ethnic group; and it may well be true that a friend of her friend is an Asian who is abusive to his wife. However, there was no reason to introduce ethnicity into the discussion. Her mention of Asian men suggested that somehow their being Asian was relevant to the abusiveness. So, just as in Haslanger's example, there would seem to be the problematic implicature that being Asian is relevant to the dangerous property being attributed. (Saul 2017, 6)

Saul (2017) argues that the *relevance implicature* and *psychological essenitalizing* that can lead to pernicious effects are not characteristical only for generics but are rather a more general phenomena. Instead, she argues, it is the occurrence of the social group terms, regardless of the type of the statement they occur in, that has essentializing effect. As she puts it, we "will remain just as likely to make these pernicious inferences about social groups, whether or not we hear the groups described with generic

language" (Saul 2017, 9). Furthermore, she stresses that it is "a red herring to focus on use of generics—the problem is with labels for social groups" (Saul 2017, 11). In other words, Saul (2017) argues, it is not a generic proposition *per se* that is the source of pernicious effects, but any mention of racial, ethnic and other groups regardless of whether occurring *within* or *outside* of a generic proposition.

I agree with Saul (2017) that psychological essentializing *per se* is not the main source of the effects of generics. Furthermore, I also agree with her that false conversational implicatures are not the source of effects of generics. Moreover, I believe that Saul's (2017) point that the kind of pernicious effects connected to psychological essentializing stemming from generics can also stem from other linguistic expressions is an important one. The importance of her argument (that the scope of the effects that psychological essentializing and relevance implicatures can contribute to is not limited to generics only) calls for broadening the scope of amelioration of these effects accordingly.

However, I also believe that we should not entirely shift our focus away from generics generic propositions either. I disagree with Saul that there is nothing special about the generic proposition that can contribute to the pernicious effects more directly. Even though, as Saul (2017) suggests, it may be true that psychological essentializing occurs in non-generic statements too, psychological research suggests that when psychological essentializing occurs in the striking property generics about social groups, the assertion rates of these generics are *higher*. And, as I argue in Section 6.3.5, *asserting generics* contributes to forming *generic judgments* which, in result, can lead to higher rates of pernicious effects, too. However, as already pointed out, psychological essentializing is not *necessary* for the effects of generics because the pernicious effects of generics can occur without psychological essentializing, whereas psychological essentializing can only contribute to higher acceptance rate of these generics.

6.3.4 Ritchie

Ritchie (2019) agrees that certain generics can be pernicious. However, she does not seem to commit to any particular view of the source of their perniciousness apart from acknowledging several authors on the topic of perniciousness such as Haslanger 2011; and Langton, Haslanger, and Anderson 2012. More importantly, however, Ritchie (2019), similarly to Saul (2017), argues that certain generics can also serve beneficial purposes. In particular, Ritchie (2019) sees *accuracy* as one of the main benefits of social generics, namely she argues that generics "can more accurately describe systematic patterns of violence and discrimination than explicitly quantified claims" (Ritchie 2019, 34).[199] In more detail, she evaluates 'descriptively accurate generic generalizations' such as "Women are expected to want

[199] "… because true generics, unlike true quantified statements, require more than mere accidental correlations. The truth of generics requires regularities or lawlike pattern. The patterns of violence and discrimination members of oppressed groups face are systematic and not accidental. When describing structural oppression, its systematicity is a core feature that ought to be captured; generics are the best tool at our disposal for doing so." (Ritchie 2019, 34)

children" as being more accurate in describing social reality than their overtly qualified counterparts such as "Some/All/Many women are expected to want children" (see Ritchie 2019, 36). This is so, she argues, because: "Existentials are made true too easily to accurately describe the systematicity and pervasiveness of social norms and patterns of discrimination," "universally quantified claims are too strong," whereas "proportional quantifiers [such as *many*] also fail to capture the generality and systematicity that generics do" (Ritchie 2019, 36).

However, it is worth noticing that Ritchie (2019) connects the above mentioned usefulness of certain generics primarily with their *descriptive accuracy* or *truth*, namely she presupposes that: "A semantic feature of generics provides additional justification for why some social generics might be effective tools for social justice" (Ritchie 2019, 36). She stresses that the main point of her paper is "that generics are able to describe the structural nature of oppression in a way that overtly quantified statements do not. This is because the truth of generics require more than accidental satisfaction and structural oppression involves more than social groups accidentally facing similar forms of discrimination" (Ritchie 2019, 37).

I believe that Ritchie is right to stress the beneficial side of certain generics. However, the scope of Ritchie's (2019) contribution to the debate that evaluates the effects of generics is rather restrictive. She focuses mainly on a particular segment of beneficial generics, namely on those generics such as "Blacks face economic, legal, and social discrimination" that seem beneficial in fighting social injustice, while also being *descriptively accurate* and *true* (see Ritchie 2019, 36). In Section 6.4.4, I challenge her claim that the role of true generics is crucial in social justice projects. Also, for why a semantic feature of generics may not be particularly relevant for the effects of generics see e.g. Section 6.3.1, Section 6.3.2, Section 6.3.5, and Section 6.4.5.3.

6.3.5 Pernicious Generic Judgments

As my reconstruction of the debate about the pernicious sources of generics shows, the main proponents of the current debate about pernicious generics agree that *generic propositions* (see Haslanger 2011; Leslie 2017; Ritchie 2019) or (at least) *terms occurring in generic propositions* (see Saul 2017) can affect speakers' reasoning and inferences which may, in result, lead to different (pernicious/beneficial/neutralizing) effects. My claim is that it is *judgments about generic propositions* that play the fundamental role in bringing about the effects of generics, whereas psychological essentializing, true or false predication, conversational or relevance implicatures can only play indirect role, by affecting or stemming from generic judgments.

In Chapter 5 (see Section 5.4.1.4), I defined generic judgments as *mental acts* that include states such as desires, attitudes of approval and disapproval, unreflective generic tendencies such as the tendency

to experience *property P being generically predicated to a given kind K,* and consciously or unconsciously held generic judgments that *property P is being generically predicated to a given kind K*.

Generic judgments play the fundamental role in the effects of generics since they guide our actions more directly than *predication*, *implicatures* and *psychological essentializing*. They can affect *reasoning* and *inferences* that can either itself have (pernicious/beneficial/neutral) effects or can lead to them. As research suggests (see Khemlani, Leslie, and Glucksberg 2012), instead of relying on the beliefs about the prevalence of the property, speakers rely on their background *generic judgments* when making inferences about individual members of a kind. As Leslie and Lerner (2016) put it:

> … accepting a generic increases people's tendency to believe a given member of the kind will have the property in question, over and above their beliefs about the prevalence of the property (Khemlani, Leslie, and Glucksberg 2012). (…) These results suggest that people may treat generics as being inferentially quite powerful, in a way that cannot be reduced to their background beliefs about the prevalence of the property. (Leslie and Lerner 2016)

It is, however, irrelevant whether *judgments about generic propositions* are generated through false or true predication, false conversational implicatures or false relevance implicatures, and with or without the help of psychological essentializing. It is irrelevant whether what is being *predicated* is actually true or false given that our judgments are not directly sensitive to the truth or falsity of a generic proposition and can, thus, be fallible. It is irrelevant whether what is being *implicated* is true or false because both false and true implicatures can affect judgments about generic propositions that can lead to effects of generics. Finally, generic judgments can occur independently of *psychological essentializing*. As already mentioned, psychological essenitalizing is not necessary for the effects but may instead *facilitate* the effects by *increasing* the assertion rate of certain generic propositions which can then result in pernicious/neutral/beneficial generics judgments (see Section 6.3.1). As an example of a pernicious generic statement that does not include psychological essentializing, consider, for instance, Ritchie's (2019) example "Blacks face economic, legal, and social discrimination."

6.4 Strategic Planning

In this section, after discussing some of the current strategies (Sections 6.4.1.–6.4.4), I put forward a novel, alternative strategy for changing the pernicious effects of generics (Section 6.4.5).

6.4.1 Leslie

Leslie's (2017) remedy for the pernicious generics is to alter our linguistic practice on the basis of the empirical evidence which "suggests that the use of labels and generics contributes to essentialization, and so the converse may also hold: reducing the use of labels and generics for racial, ethnic, and religious groups may reduce the extent to which children grow up essentializing these

groups" (Leslie 2017, 420). In particular, Leslie (2017) suggests rephrasing generics that include labeling people as a 'Muslim' or a 'Black' or an 'African American' with a description 'a person who follows Islam' or 'a person with darker skin,' respectively. In this way, she argues, we emphasize "that person is the relevant kind sortal" and that following Islam or having darker skin "is a particular property that the individual happens to possess" (see Leslie 2017, 420). Furthermore, Leslie (2017) speculates that her proposal will have some immediate benefits because:

> … findings by Carnaghi and colleagues suggest that hearing a member of a familiar social kind described by an adjective rather than a noun can reduce the extent to which adults expect the individual to conform to a stereotype. It is possible, though, that the real benefits would extend beyond the alteration of the attitudes of adults; the really intriguing possibility would be to decrease the extent to which children in our society grow up essentializing social groups. (Leslie 2017, 420)

As a corollary, Leslie's (2017, 420) proposal invites a large-scale linguistic revisionism, and, ultimately, an *abandonment* of a certain subclass of linguistic expressions, i.e. generics about social groups. As she puts it: "Eschewing reference to social kinds, at least by way of labels and generics, may be a direct way to reduce the extent to which we unconsciously teach our children to essentialize" (Leslie 2017, 421).

Given her utilization of the empirical data, Leslie's proposal has been recognized by Cappelen (2018, 29) as pertaining to the *domain-specific project of conceptual engineering* (see Chapter 2, Section 2.1.8). I agree that the main strength of Leslie's proposal is her utilization of the empirical data in engineering generics. In my approach to engineering generics, I also endorse the utilization of empirical data, both during the deliberation process of the strategic planning stages (see Section 6.4.5) and by means of informing the implementation stage (see Section 6.5). However, below I offer two reasons why we should doubt Leslie's particular proposal.

First. Leslie's (2017) proposal is not an optimal strategy because it is both *self-undermining* as well as *hard to implement*. In particular, Leslie utilizes empirical data which shows that we tend to remember quantified statements as generics in order to support her claim that generics 'give voice to' a default mode of generalization for our minds. The hypothesis Leslie (2007, 2008) uses to explain the claim that generics are cognitively more fundamental generalizations as opposed to the quantified generalizations is that the latter are more taxing and sophisticated, and, thus, need an explicit instruction such as "all" or "some." As Leslie (2014) puts it:

> If the cognitive system has a basic, default way of forming general judgments then it may sometimes fall back on this means of generalizing when asked to process a more taxing and sophisticated generalization. On this hypothesis, quantified statements are precisely this, namely, more taxing and sophisticated generalizations. Thus, if generics do articulate cognitively fundamental generalizations, it is easy to understand why both adults and children show a tendency to "default" to the generic when asked to consider quantified statements. (Leslie 2014,

210)

However, if the above claims she endorses hold, then her proposal which consists in abandoning generics by turning them into quantified statements is going to be futile because our minds will try to find a way back to this default mode of generalizing, i.e. to express the quantified statements as generics. In that sense, generics seem inevitable. In other words, the same data Leslie uses to explain (at least partly) why generics can be pernicious, can be also used to predict the unsuccessfulness of her strategy.

Saul (2017) raises a similar worry. She argues that merely changing things at the level of generic language will not help with getting rid of the pernicious effects of generics: "If Leslie is right that our minds tend to leap by default from single dangerous instances to very general beliefs, then we should expect this to happen regardless of the words that are used" (Saul 2017, 8).[200] In other words, if striking property generics cannot be avoided because they 'give voice to' our default mode of generalization, then even though "it might be true that when we initially replace 'Muslim' with 'person who follows Islam' (a suggestion that she makes), we'll be slower to ascribe an essence. But soon that phrase will simply be a label, and function as one" (Saul 2017, 14). She argues that given that there is no technical obstacle, we would soon start constructing generic propositions like "People who follow Islam are terrorists."[201]

Second. If Leslie's (2017) is right, then we should avoid all generics about social groups.[202] However, rephrasing and eradication of our use of generics may be *too radical* or even *unnecessary*, especially if applied to all generics about social groups. I believe that, as Saul (2017) has argued, we should not abandon generics as long as we have a good reason not to see them as peculiarly destructive (or at least more destructive than many other similar constructions). Avoiding generics would be an error since they, in fact, serve a valuable purpose and can be useful in the cause of social justice. I also agree with Saul that psychological essentializing does not need to be seen only as a 'cognitive mistake' we want to get rid of but can instead be seen as a *cognitive advantage*. I believe that we could utilize the phenomenon of psychological essentializing connected to generics instead of avoiding it. For instance, psychological essentializing of generics can be used for introducing beneficial generics such as "Girls

[200] A similarly worry may apply to Langton, Haslanger, and Anderson's (2012, 765), suggestion that in the case of statistically true generics in order to avoid essentializing we should use covertly quantified statements.

[201] Saul supports her claim in the following way: "It is worth reflecting also on how notably unsuccessful it was to replace the noun 'moron' with the descriptive phrase 'mentally retarded person'. The more recent terms 'special needs' and 'person with special needs' also provide a revealing case study. Indeed, the drive to label groups with noun phrases has led the noun phrase 'special needs' to be used as an adjective in 'special needs children'. Even when it's (initially) ungrammatical, we will find a way to form the easy noun phrases that facilitate essentialising. Reflecting on cases like these should give one pause about the efficacy of attempting to reduce essentialising through this sort of linguistic reform" (Saul 2017, 14).

[202] Apart from Leslie 2017, the proposals of e.g. Haslanger 2011; Langton, Haslanger, and Anderson 2012; Wodak, Leslie, and Rhodes 2015; Wodak and Leslie 2017 can also be seen as resulting in avoidance of racial and gender generics (see Ritchie 2019).

are good at math" (see Saul 2017; Ritchie 2019; Section 6.7).

6.4.2 Haslanger

Haslanger (2011) argues that even if pernicious generics were true, they should be *denied* because they carry *false implicatures*. Because she wants to remain neutral with respect to the truth-conditions of such generics (see Section 6.3.2), she does not want to endorse a *descriptive negation*, e.g. to claim that "Blacks are violent" is *false*. Instead, she proposes a remedy at the *metalinguistic level* by utilizing *metalinguistic negation*. By using metalinguistic negation, Haslanger believes that she can oppose the ideology behind such generics and avoid the talk about their truth-conditions.

As opposed to a descriptive negation, *metalinguistic negation*[203] (introduced by Horn 1989) has been used to demarcate pragmatic ambiguity. For example, consider the following cases where one speaker objects to the previous speaker's utterance on the grounds such as phonetic or grammatical form, register, or associated presuppositions or implicatures:

a. Around here we don't LIKE coffee—we LOVE it.

b. She doesn't sell INsurance—she sells inSURance.

c. It's not stewed bunny, honey, it's *civet de lapin*.

d. I'm not HIS brother—he's MY brother!

e. Mozart's sonatas were for piano and violin, not for violin and piano.[204]

Haslanger (2011), however, does not offer any details of how the metalinguistic negation should be formulated in the case of generics. Saul (2017, 5) suggested that the metalinguistic negation solution that Haslanger (2011) endorses for a generic "Blacks are violent" may look like this: "No, blacks aren't violent! *People* are violent when they're placed in oppressive circumstances." Or: "No, blacks aren't violent! They don't share some common nature, but in fact are as diverse as any other group of people."

I believe that Haslanger (2011) rightly detects that the effects of generics are connected to the *pragmatics* of generics. Furthermore, for the reasons explained in Section 6.4.5, I believe that Haslanger (2011) is right to suggest not to engineer the effects of generics at the semantic level. However, given these commitments, I believe that utilizing a metalinguistic negation is not a promising strategy either. Contrary to Saul (2017), who does not see Haslanger's suggestion to utilize metalinguistic negation as a

[203] See also *echoic negation* (Carston 1996).

[204] See Horn and H. Wansing 2020.

real worry,[205] I believe that utilizing a metalinguistic negation may leave one with a worry of having to commit to the truth value of a generic proposition. In particular, Horn (1989) introduced *metalinguistic negation* for a *denial of true statements* with false implicatures. Given that, it may be argued that Saul's example of *metalinguistic negation*[206] which includes "No, blacks aren't violent!" requires from the speaker both to *think about* as well as to *commit* to a truth value of a particular generic statement. If so, this seems to clash with Haslanger's (2011) commitment to stay neutral on the truth value of generics, and her claim that we are not good at thinking about their truth-conditions.

A defender of Haslanger's proposal might respond that the worry that I just raised is an artefact of the particular example chosen by Saul. Instead, a defender might suggest that the following exchange would not require one to obviously commit to the truth value of a generic proposition: "Hang on a minute! *People* are violent when faced with daily oppression."[207] However, I believe that, in this example, "Hang on a minute!" reads as an *invitation to belief refrainment* that "Blacks are violent," whereas the paradigmatic cases of the "metalinguistic negation" and "echoic negation" include explicit negation. Furthermore, looking back at Saul's proposal, one could perhaps read the part of her proposal "No, Blacks are not violent" as a *belief refusal* that "Blacks are violent" rather than a claim about the truth value of "Blacks are violent." If so, both of these strategies, i.e. belief refrainment and belief refusal, are compatible with my proposal which invites one to change one's doxastic propositional attitude towards the generic proposition in question (see Section 6.4.5.4).

6.4.3 Saul

In a nutshell, Saul's (2017) proposal is that "we shouldn't try to avoid generics. Instead, we need to *get better at talking and thinking about them*. We need to press people to spell out their evidence for their generic claims and to reflect on what that evidence really does or doesn't warrant" (Saul 2017, 14; italics mine). She illustrates her strategy by offering the following example:

> Suppose that A asserts 'Muslims are terrorists'. B could ask any of the following.
> - A few Muslims are terrorists. But so are a few Christians. Remember Timothy McVeigh?
> - Why do you think that? How many terrorist Muslims do you know about out of the more than a billion in the world?
> - Do you have any evidence that being Muslim makes you more likely to be a terrorist? (Saul 2017, 14–15)

[205] "… a key part of her [Haslanger's] point is that we're not very good at thinking about the truth conditions for generics. She herself is not sure what they are, and—perhaps even more importantly—we are sure to encounter many people who are convinced that (5) is true. Given all this, it may well be a good idea to use metalinguistic negation in order to block the addition of the implicated claim to the common ground. This way we can protest the ideology, while sidestepping a discussion of the truth conditions of generics." (Saul 2017, 5)

[206] Similar point *might* apply to approaches that would rely on *metalinguistic negotiation* (see Ludlow 2014 and Plunkett and Sundell 2013) or *communicative disruptions* (see Sterken 2020).

[207] Thanks to Mark Pinder for suggesting this example.

Saul admits that her strategy can be seen as similar to Haslanger's in that it permits the use of metalinguistic negation in the following way: "No, Muslims are not terrorists. A few terrorists who are members of the group aren't enough to justify what you said" (see Saul 2017, 14–15). However, she also stresses that, contrary to Haslanger's, her "strategy is not confined to metalinguistic negation, and it's not motivated by the thought that there is an implicature about natures in need of cancellation" (Saul 2017, 14–15).

Furthermore, Saul stresses that it would be a mistake to limit our focus only on generics. Instead, she believes, we should pay attention and apply this strategy to all mentions of "race, gender and the like" since "social kind membership is often mentioned when it's not relevant both in generics and non-generic cases" (Saul 2017, 15). She points out that ideology such as racist is often transmitted without explicit mention of race (*linguistically*: e.g. dog-whistle terms (see Saul 2018, Stanley 2015); or *non-linguistically*: simply by living in a world structured by residential, occupational and educational segregation (see Anderson 2010).

Ultimately, we can understand Saul's proposal as *abolition of unwarranted social group terms upon reflection* in certain contexts rather than across the board. Most importantly, Saul insists, we should do regardlessly of whether social group terms occur in generic or some other propositions. As she puts it: "We need to call attention to them and ask whether they're relevant, demand evidence for claims, question inferences and so on. We definitely should do these things when generics are used. But we should also do them when generics are not used" (Saul 2017, 16–17).

I welcome Saul's (2017) approach that utilizes *rational reflection* as a tool to contribute to the amelioration of pernicious effects. However, even though rational reflection can be beneficial, I believe that expecting that the speakers themselves always spell out the evidence could prove *hard to implement* (especially among the children) and *might not (always) be necessary*. Moreover, given that speakers are not *ideal reasoners*, and given the difficulty of understanding and knowing the truth-conditions and the truth value of generics, it might be *too demanding on the speakers* to (always) use rational reflection in the way Saul (2017) proposes. Finally, the application of this strategy might (at least sometimes) become a moot point since the speakers could fall back on a generic form and claim that by saying e.g. "Muslims are terrorists" they do not mean that *all Muslims* are terrorists but 'simply' *Muslims* are terrorists. When pressed for further explanation, they may refuse to think or spell out further what they mean by their claim or how many Muslims have to be terrorists in order for what they are claiming to be true. And, perhaps, even rightly so given that the theorists too struggle with giving a unified answer to this question. Relatedly, one may argue that insisting on giving further reasons will *change the subject of conversation*. To conclude, I believe that Saul's strategy is useful and should be employed (when possible), but is also quite demanding and so we should look for additional strategies to complement it.

6.4.4 Ritchie

The gist of Ritchie's (2019) proposal is that *we should avoid some but not all* of racial and gender generics. As an exception to avoidance, she considers those generics which can be used in social justice projects since:

> Social justice projects often have both descriptive and prescriptive aims. We need an accurate description of where we are now and how we got here, the thought goes, in order to determine what strategies can best mitigate oppression. If some generic claims about race, gender, and other social categories are more precise descriptions of social reality than explicitly quantified statements, the descriptive component of a social justice project would be better served through the use of generic claims. (Ritchie 2019, 37)

Apart from the purposes of social justice projects (which, according to her, require the use of *true* generics because of their *accuracy*), Ritchie (2019) briefly admits the possibility of a beneficial utilization of social generics even if they were *false*. She evokes Saul's (2017) suggestion that generics such as "Girls play football" (which they both take to be false) might be "very important weapons in our anti-prejudice arsenal" (Saul 2017, 13). Ritchie, thus, considers the possibility that "even if it is a false generic generalization, there may be good reasons to *assert* it in certain political contexts" (Ritchie 2019, 38; italics mine). However, she admits that there are limits to the scope of her main proposal since: "The claim argued for here is more modest: the ability for generic generalizations to accurately describe systematic patterns of injustice is one explanation for their social political expedience" (Ritchie 2019, 38).

I agree with Ritchie (2019) that generics can be utilized for beneficial purposes. However, as she admits, the scope of her suggestion is rather limited (especially given the scope of this chapter). Her focus is primarily on utilizing *true* generics for social justice projects. More problematically, her main suggestion would, ultimately, require one to comply with the *truth assertion norm*, i.e. a norm that, in its simplest formulation, says that one should assert the truth. Moreover, her relying on the fact that the truth of generics requires regularities or lawlike pattern may turn out to be problematic since the truth value of generics may be a result of something that does not require regularities or lawlike patterns, such as e.g. arbitrary reference (see Fuš manuscript b).[208] Furthermore, even if the truth of generics does require regularities or lawlike patterns, we might not be in a position to know whether a particular generic is true (see Haslanger 2011; Fuš manuscript b). Finally, given that Ritchie (2019) still allows for a partial prohibition of certain generics, this means that the objections similar to the ones raised against

[208] In Fuš (manuscript b), I propose a view I dub *Arbitrary Reference Account of Generics,* an epistemicist view about the semantics of generics. This view builds on Breckenridge and Magidor's (2012) epistemicist view of Arbitrary Reference, and rests on Kearns and Magidor's (2012) defense of the metasemantic view of *Semantic Sovereignty*. In particular, I defend a thesis that generic statements have truth values yet we do not and cannot know whether generic statements are true or false. Everything is in order with classical logic, i.e. there is no indeterminacy (i.e. there is a sharp cut-off point of the domain generic quantifier ranges over) whether generics are true or false, only epistemic uncertainty (i.e. we cannot possibly discover it nor know how it works). Instead, we are deeply ignorant how metasemantics of generics work and, consequently, we are deeply ignorant about the semantic content and truth value of generics. For example, in the case of generic expressions such as "Muslims are terrorists" the generic is either true or false, yet we do not and cannot know whether is true or false.

Leslie (2017) (see Section 6.4.1) translate to Ritchie's (2019) view.

6.4.5 Engineering Generic Judgments

In the evaluation section, I have argued that the pernicious effects of generics stem most fundamentally from our *generic judgments*. Given that, I want to propose *engineering generic judgments*. There may be different ways to engineer generic judgments, however. In this section, I outline the *Speech Act Approach* as one promising way to change the (pernicious) effects of generic judgments.[209] In a nutshell, the Speech Act Approach achieves a change in doxastic (i.e. belief, disbelief, belief refrainment) propositional attitudes towards generics by utilizing speech acts.

6.4.5.1 Generic Judgments and Propositional Attitudes Towards Generics

In the literature about *judgments*, it is often assumed that, even though not all, at least some judgments can have propositional or conceptual content which is 'language-like' or 'discursive.' Having this representational format is often explained as having inner 'language of thought' (LOT) (see Fodor 2008), and being dependent on the mastery of a public language (see Brandom 1994; McDowell 2009).

Propositional attitudes are mental states or cognitive reactions that an agent can have towards propositions. Propositional attitudes include *beliefs*, *desires*, *hopes*, *wishes*, *intentions*, *fears*, *doubts*, etc. Some propositional attitudes have 'direction of fit,' namely, some propositional attitudes aim to reflect or to represent the world, another to influence or to change it. The content of propositional attitudes can be *true* or *false from the perspective of an agent*, and an agent can have different propositional attitudes towards the same proposition. For instance, in the case of propositional attitudes towards generic propositions, an agent can *believe* that Blacks are violent, an agent can *fear* that Blacks are violent, an agent can *wish* that Blacks are violent, an agent can *doubt* that Blacks are violent, etc.

My proposal relies on the assumption that *speakers' generic judgments* are closely connected to *speakers' propositional attitudes towards generics*.[210] The assumption takes it that there is a correlation between *generic judgments* and *propositional attitudes towards generics*, i.e. that, even though not all, at least some *generic judgments* can result in *propositional attitudes towards generics* as well as *vice versa*, namely that, even though not all, at least some *propositional attitudes towards generics* can result in *generic judgments*. For instance, if one *judges* that Muslims are terrorists then this may result in one forming a

[209] While deliberating whether and which of the general philosophical engineering strategies would be the most plausible, it is worth to keep in mind that the main target of this chapter are the (pernicious) effects of generic propositions. Furthermore, what one takes as a source of these effects can (at least partly) determine what kind of strategies will be available. There may also be more strategies that are available (pluralism about strategies) and there may be more strategies that are implementable (pluralism about implementation). Finally, whereas some strategies might be theoretically more optimal to change the effects of generics, their implementation might be suboptimal.

[210] Van der Schaar (2013, 101) claims that early Stout, Russell and Moore shared the following thesis: "The object of judgement has a propositional structure, and may function as object of other propositional attitudes" (van der Schaar 2013, 101; italics mine).

propositional attitude of *belief* that the generic proposition "Muslims are terrorists" is true. And *vice versa*. If one forms a propositional attitude of *belief* that the generic proposition "Women are submissive" is true then this may result in the mental act of *judging* that Women are submissive.

Given this assumption, my hypothesis is that changing propositional attitudes towards generics could contribute to changing generic judgments and, thus, contribute to *changing the effects* of generic judgments. With this in mind, this chapter has certain limitations. First, I focus primarily on *engineering* (pernicious/neutral/beneficial) effects of those generic judgments that can result in the agent forming *propositional attitudes towards generic propositions*, and *vice versa*. Furthermore, since there may also be non-linguistic ways to achieve generic judgments, a particular engineering strategy that I propose may not be *sufficient* nor *necessary* to change all generic judgments but rather those that result *in* and *from* the speakers forming attitudes towards generic propositions.

In the next sections, I focus on illustrating the prospects of engineering of the *classificatory doxastic attitudes* towards generic proposition: *generic belief* (believing a generic proposition to be true), *generic disbelief* (believing a generic proposition to be false), and *generic (dis)belief refrainment*[211] (refraining from believing that generic proposition is true or false).[212] In Section 6.4.5.2, I outline the Speech Act Approach to engineering *classificatory doxastic attitudes* towards generics which may, consequently, lead to *engineering of generic judgments*. In Section 6.4.5.3, I suggest some possible strategies.

6.4.5.2 The Speech Act Approach

The Speech Act Approach utilizes *speech acts* in order to achieve change in audience's *generic belief*, *generic disbelief*, or *generic belief refrainment*. To get the Speech Act Approach off the ground, the *implementator* should have the *authority*[213] over the audience in such a way that the audience *recognizes* and *accepts* the implementator's *reflexive communicative intention* to assert, refuse or refrain from asserting (/refusing) that G.

In pragmatics, reflexive communicative intention is an intention to produce a belief or other propositional attitude by means (at least in part) of the audience's recognition of this very intension to produce a belief or other propositional attitude (see Grice 1957, Strawson 1964, Searle 1969). The reflexive communicative intention is fulfilled when the audience recognizes the speaker's (in this case,

---

[211] Despite receiving little attention in the recent debates in epistemology, *belief refrainment* has its roots in the Stoics philosophy, and has important place in Sextus Empiricus' Pyrrhonian skeptic, as well as in Descartes' epistemology. Belief refrainment is a propositional attitude that often occurs (and rationally so) when, for example, a certain subject matter has not yet been considered, or it needs to be considered further, or it cannot be considered. As one of the well known instances of belief refraining consider a legal principle called '*The presumption of innocence*' which says that one is considered innocent until proven guilty. For example, we should refrain from judgment about whether e.g. Tom Hagen (the prime suspect in the murder investigation) murdered his wife.

[212] The object of a *generic belief*, *generic disbelief* and *generic belief refrainment* are *generic propositions* which are expressed by generic sentences, e.g. a generic proposition "Blacks are violent" expressed by a generic sentence 'Blacks are violent.'

[213] The authority of the implementator should be justified and governed by the underlying engineering project.

implementator's) intention.[214] In the case of generics, for instance, psychological research suggest that children: "upon hearing "Zarpies are good at baking pizzas", they assumed not only that a new Zarpie would be good at this, but also that a new gorp would not. Children only made these inferences when generics were said by *knowledgeable speakers*, suggesting that children reason about a speaker's intended meaning to draw these inferences" (Moty and Rhodes forthcoming; italics mine).

If the implementation is successful, namely if the implementator achieves to make the audience to *recognize* and *accept* what the implementator is asserting, refusing, or refraining from asserting(/refusing), this would put the receptive audience in a position to *adopt* the implementator's underlying propositional attitude of belief, disbelief, or (dis)belief refrainment that G, which would result in a *change* in the audience's own belief, disbelief, or (dis)belief refrainment that G. Finally, this would result in a change in the audience's generic judgment.

Moreover, in the pragmatic literature it is often taken for granted that when one believes, disbelieves, or refrains from (dis)belief that P and is in conditions that allow for sincere expression of that belief, disbelief, or (dis)belief refrainment, one will typically assert, refuse, or refrain from asserting (refusing) that P. Given that, a successful implementation would not only change the audiences' generic judgments but would also *cause the audience to* assert, refuse, or refrain from asserting(/refusing) the intended G which may aid the implementation of the intended propositional attitude towards G.

It is, however, worth mentioning that there may be certain constraints that can make the audience be more or less receptive. As Pinder (2019) points out: "our expectations of the audience may be affected by the audience's linguistic sophistication, the complexity of the definition, the linguistic context, the political and social climate, and so on." (Pinder 2019, 13) Nevertheless, they could also be overcome. One can apply a similar procedure as the one that Pinder (2019) does when he suggests changing the speaker-meaning: "But if the speaker is explicit about her intentions then, in an appropriate context, with an appropriate audience, speaker-meaning can deviate from semantic-meaning in new and substantial ways, and the audience will be able to grasp that speaker-meaning" (Pinder 2019, 13).

Thus, if the implementator has sufficient authority over the appropriate audience in an appropriate context, the receptive audience should recognize and accept the implementator's reflexive communicative intention to assert[215] that e.g. "Girls are tough," or refuse that e.g. "Blacks are violent," or refrain from asserting that e.g. "Martians are evil." This would put the audience in a position to adopt the implementator's underlying propositional attitude to *believe* that e.g. Girls are tough, to

---

[214] As Searle put it: "we achieve what we try to do by getting our audience to recognize what we try to do" (Searle 1969, 47).

[215] Assertions often express beliefs but are not the only way to express belief since there are linguistic acts (e.g. such as invitation) that can express belief but are not assertions (see MacFarlane 2009). This could mean that we could change believes by other means. However, in this chapter, I focus on assertion, refusal, and refraining from assertion/refusal of a generic proposition as speech acts which could lead to a belief, disbelief, or refrainment from (dis)belief.

disbelieve that e.g. Blacks are violent, or to *refrain from belief* that e.g. Martians are evil. Consequently, this would also cause the audience to *assert* that e.g. "Girls are tough," or *refuse* that e.g. "Blacks are violent," or *refrain from asserting* that e.g. "Martians are evil" which may aid the implementation of the intended propositional attitude towards the respective generic propositions.

6.4.5.3 Strategies

Below are some of the more particular available strategies that the implementor of the Speech Act Approach could utilize. I categorized the strategies by taking into consideration three general modes of philosophical engineering: (a) *reengineering*, (b) *replacement* and (c) de novo *engineering* (see Chapter 2).[216]

(a) Reengineering

(i) changing *asserting* a certain generic proposition (e.g. Blacks are violent) to *refusing* the same generic proposition instead;

(ii) changing *refusing* a certain generic proposition (e.g. Girls are good at math) to *asserting* the same generic proposition instead;

(iii) changing *asserting* a certain generic proposition (e.g. Women are submissive) to *refraining from asserting* the same generic proposition instead;

(iv) changing *refusing* a certain generic proposition (e.g. Boys love pink) to *refraining from refusing* the same generic proposition instead;

(v) changing *refraining from refusing* a certain generic proposition (e.g. Muslims are terrorists) to *refusing* the same generic proposition instead;

(vi) changing *refraining from* asserting a certain generic proposition (e.g. Girls are tough) to *asserting* the same generic proposition instead.

(b) Replacement

(i) countering *asserting* a certain generic proposition (e.g. Latinos are lazy) with *asserting* a newly introduced generic proposition in the vicinity (e.g. Latinos are sometimes lazy);

[216] The examples utilized here are for illustrative purposes only.

(ii) countering *asserting* a certain generic proposition (e.g. Boys will be boys) with *refusing* a newly introduced generic proposition in the vicinity (e.g. Boys will always be boys);

(iii) countering *asserting* a certain generic proposition (e.g. Blacks are violent) with *refraining from asserting* a newly introduced generic in the vicinity (e.g. Blacks cannot be violent);

(iv) countering *refusing* a certain generic proposition (e.g. Girls play football) with *refusing* a newly introduced generic proposition in the vicinity (e.g. Girls cannot play football);

(v) countering *refusing* a certain generic proposition (e.g. Girls are tough) with *asserting* a newly introduced generic proposition in the vicinity (e.g. Girls can be tough);

(vi) countering *refusing* a certain generic proposition (e.g. Girls play football) and with *refraining from refusing* a newly introduced generic in the vicinity (e.g. Girls cannot play football);

(vii) countering *refraining asserting/refusing* a certain generic proposition (e.g. Boys don't cry) with *asserting* a newly introduced generic proposition in the vicinity (e.g. Boys can cry);

(viii) countering *refraining from asserting/refusing* (e.g. Boys don't cry) a certain generic proposition with *refusing* a newly introduced generic proposition in the vicinity (e.g. Boys never cry);

(ix) countering *refraining from asserting/refusing* a certain generic proposition (Boys don't cry) with *refraining from asserting/refusing* a newly introduced generic in the vicinity (e.g. Boys cannot cry);

(c) De novo *engineering*

(i) introducing an *assertion* of a generic proposition (e.g. Black lives matter; Blue lives matter);

(ii) introducing a *refusal* of a generic proposition (e.g. Martians are killers; People without alibi are always guilty);

(iii) introducing *refrainment from assertion/refusal* of a generic proposition (e.g. White silence is violence).

6.5 Implementation

For the purposes of this chapter, I largely remain focused on a domain-specific *theorizing* and, thus, do not offer any particular implementation solution for any particular instance of a generic judgment. Instead, I offer some general considerations for implementing the Speech Act Approach to engineering generic judgments.

6.5.1 The Role of Empirical Research

As illustrated in the previous sections, empirical data on generics has already given us a valuable insight into psychological and evolutionary pressures that may have shaped our generic judgments and our propositional attitudes towards generic propositions. I believe that the insights from empirical data should play an important role in engineering effects of generic judgments. For example, empirical evidence may inform us about the motives and the underlying mechanisms present in formation of certain propositional attitudes towards generic propositions. However, we *should* avoid *Hume's (is-ought) law,* namely we should not conclude from the fact that we *currently have* certain propositional attitudes towards generic propositions (e.g. that we *currently believe* certain generics) that we *ought to have* these propositional attitudes towards generic propositions (e.g. that we *ought to believe* certain generics). Instead, empirical research should be in the service of informing and facilitating the change of propositional attitudes towards generics in order to achieve desired effects.[217] Furthermore, empirical work can help us determine whether our proposal has been successful. As Nado 2020, points out:

> We don't have direct epistemic access to the classificatory dispositions of our fellows, but their behaviour and verbal reports can give us a pretty good indication of whether the desired change has taken effect. (…) For instance, we can run a poll to see how many Americans believe that transgender women count as women. This won't tell us whether the meaning of 'woman' has changed; but it will tell us quite a lot about the classificatory procedures Americans use. (Nado 2020, 24)

A similar strategy that Nado propose for tracking the change in classificatory dispositions can be utilized to determine the change of propositional attitudes.

[217] E.g. psychologists have been conducting studies in which they have shown how even a small-scale manipulation of certain conditions can bring upon change in the participant's acceptance, rejection or refrainment of certain generics (see Moty and Rhodes forthcoming.).

6.5.2 Natural Selection and Rational Reflection

Empirical research suggests that *evolutionary forces* have been shaping our generic judgments significantly. However, there is some evidence that another, *compatible*, force *has* also *shaped* and, I believe, *should be shaping* our generic judgments and our propositional attitudes towards generics—*the force of rational reflection*.

In Chapter 5, I endorsed a *Darwinian assumption about generic judgments* which says that our *generic* judgments have been indirectly yet significantly influenced by natural selection. I argued that *the adaptive link account* offers a better scientific explanation for why we have certain generic judgments than its *competitor.*[218] Moreover, I argued that the *influence of rational reflection* on our acceptance/refrainment of generic judgments is compatible with the adaptive link account (see Chapter 5, Section 5.5.2).[219] One can, for instance,[220] imagine that, given enough time and change in our natural and social environment, evolutionary forces can influence at least some of our propositional attitudes towards some of our current generic judgments to change or disappear, or new generic judgments to get introduced. For example, imagine the plausibility of asserting "Mosquitoes carry the West Nile virus" in the total absence of an existence of mosquitoes or the West Nile virus. Or, imagine the plausibility of asserting a generic proposition such as "Black lives matter" or "White silence is violence" before passage of the Thirteenth Amendment from 1865.

Nevertheless, it is worth keeping in mind that it is a separate question what are the good reasons for intervention by rational reflection, especially after some generic judgments have already been favored by evolution.[221] Do engineers have good reasons to intervene in something that may affect our survival and reproduction and what sort of reasons would that be? For instance, upon some rational reflection, one may argue that endorsing generic judgments such as "Muslims are terrorists" would result in more social, moral and political harm than survival and reproductive good. Or, one may ask what should prevail in deliberating about asserting or refusing generic propositions such as "Sharks eat swimmers," our survival rate or the animal rights and their survival rate?

---

[218] *The Adaptive Link Account for Genericity:* tendencies to make certain kinds of generic judgments rather than others contributed to our ancestors' reproductive success not because they constituted perceptions of independent generic truths, but rather because they forged adaptive links between our ancestors' circumstances and their responses to those a, getting them to act, feel, and believe in ways that turned out to be reproductively advantageous. To elaborate: As a result of natural selection, there are in living organisms all kinds of mechanisms that serve to link an organism's circumstances with its responses in ways that tend to promote survival and reproduction.

[219] Analogously, Street argues that: "It is perfectly compatible with the adaptive link account that we come to reject some of our basic evaluative tendencies on the basis of other evaluative judgements we hold" (Street 2006, 161). Similarly, in Chapter 5, I argued that *rejecting* some of our basic generic tendencies on the basis of other generic (and perhaps other, non-generic) judgments we hold is compatible with the adaptive link account.

[220] However, the above examples are also compatible with the view that does not endorse the role of evolutionary forces or rational reflection, such as a possible view that our *generic judgments* and *propositional attitudes towards generics* can be sensitive only to factors such as *change of the state of affairs, time, context, social norms*, etc.

[221] For more about the wrong and the right kind of reasons in the context of conceptual engineering, see Scharp 2020, 409–410; Section 6.8.

6.5.3 Being Out of Control

In the context of philosophical engineering, Cappelen (2018) defends a general pessimism about 'being in control' of implementing particular engineering projects:

> Anyone who spends time thinking and talking about large-scale normative matters should do so without holding out too much hope that their talking and thinking will have significant or predictable effects on the relevant aspect of the world. (Cappelen 2018, 200)

> We can observe these defects, describe them, reflect on them, and think of ameliorative strategies. But careful thinking also reveals that such reflection is ineffective. Amelioration might happen, but if it does, it has little to do with our intentional efforts. Our intellect can diagnose itself, figure out a cure, but is impotent when it comes to doing anything. Emphasizing this highlights an important limitation on human rationality and intellect. (Cappelen 2018, 201)

Even though a number of authors has already offered a defense against this kind of pessimism when it comes to particular subkinds of philosophical engineering (see Chapter 4), here I offer a brief defense of how utilizing an intentional effort, i.e. *rational reflection* to come up and implement the intended strategies, may work in the case of generics.

Given that evolutionary forces might be *ultimately* responsible for keeping, abandoning or refraining from certain generics judgments, one may also take into consideration the impact of natural selection on rational reflection as one of the potential impediments to implementation. My response to this worry is that *rational reflection* can be seen as part of natural selection. This reply is compatible with my general view of *philosophical engineering implementation* (see Chapter 2, Section 2.3) on which *natural selection* and *rational reflection* can go hand in hand. For example, in the context of philosophical engineering, upon some *rational reflection* during the strategic planning stage, philosophical engineering strategies get proposed and dropped into the *philosophical gene pool*.[222] Even though we might not be able to control nor predict the process of natural selection, rational reflection can still influence natural selection by facilitating the process of natural selection by passing on strategies with better *fitness* (see Chapter 5, Section 5.4.1.3).

Measuring a particular implementation success is an empirical question, and we have not yet started testing nor engaging in any systematic change of generic judgments. Nevertheless, in order to provide some support that the force of rational reflection may be utilized in the case of the Speech Act Approach to engineering generic judgments, I offer some illustration of the impact that rational reflection has had on some of the universally and existentially quantified generalizations. In particular, empirical data suggests that, as opposed to generalizations via overt quantifiers such as *universally* and *existentially quantified generalizations* which are cognitively more taxing and sophisticated, *generic*

[222] Thanks to Patrick Greenough for suggesting this term.

generalizations are cognitively fundamental and automatic way of making generalizations (see Leslie 2007, 2008, 2012, 2014). Nevertheless, against the psychological and evolutionary odds, by utilizing the tools of *rational reflection*, speakers still make an effort to *formulate* universally and existentially quantified generalizations as well as to *engage in forming and changing* propositional attitudes towards them. My suggestion, similarly, operates on the assumption that against the psychological and evolutionary odds, we could and should utilize *rational reflection* to *engage in forming and changing* propositional attitudes towards generic propositions.

Finally, it is worth keeping in mind that even if *evolution* and *rational reflection* can go hand in hand, it may still take time before the *adaptive link* can be properly forged. And even if the adaptive link does not get forged, sometimes a temporal change of some of our generic judgments might count towards partial social, moral or political amelioration.

6.5.4 Engineering Holism[223]

As mentioned in the previous sections, changing effects of generic judgments may involve different kinds of engineerings. In this chapter, I proposed utilizing speech acts in order to engineer a certain subclass of *propositional attitudes towards generic propositions* as *means* of achieving another *target engineering*—e.g. *engineering generic judgments*. However, one can see engineering generic judgments as *means* to another end, e.g *social, moral or political engineering*.

Furthermore, it is not entirely excluded that e.g. metaphysical, semantic or epistemic change could contribute to a change in the rate of *assertion* or *refusal* of generic propositions (and, possibly, *vice versa*). For instance, semantic or epistemic engineering of a concept "women" might affect an assertion rate of "Women are submissive." Or, a metaphysical change such as extinction of the West Nile virus could lower the assertion rate of "Mosquitos carry the West Nile virus." Or, a change in what is perceived or known as a dangerous property could lower the assertion rate of some of the striking property generics. Or, resolving the underlying oppression, which is often thought to produce the phenomenon of *illocutionary silencing*, could enable change in propositional attitudes towards generics. In particular, getting rid of the underlying oppression might lift *illocutionary silencing* since for some individuals and some (e.g. racial, religious, and ethnic minority) groups it might be the case that *different forms of oppression*[224] hinder implementation of certain speech acts by causing their illocutions to *misfire* (e.g. consider women in pornography (see Langton 1993, MacKinnon 1992 and Hornsby and Langton 1998);

---

[223] See also Chapter 2, Section 2.3.3.

[224] "McGowan 2009 argues that some speech acts can not only cause but also constitute instances of oppression. Anderson, Haslanger and Langton (2012) and McConnell-Ginet (2012) provide overviews of research on racial, gender and related forms of oppression as they relate to speech acts." (Green 2020)

or ethnic groups during the apartheid).[225] In the case of generics, illocutionary silencing could apply in some contexts to certain individuals or groups of people (e.g. the group of Blacks) who may be able preform normal locutionary acts but not being able to preform the illocutionary acts (e.g. asserting or refusing generics such as "Blacks are violent"), thus, making their illocutionary act "inert." In these cases, getting rid of the underlying oppression might be a precondition for their illocutionary act (e.g. assertion or refusal) to be preformed.

6.5.5 Outsourcing Implementation

Given engineering holism that I endorse, the implementation of engineering generic judgments via speech acts could be outsourced to different disciplines or kinds of engineerings which may have enough authority to change audiences' propositional attitudes and speech acts. Here are some *ad hoc* candidate practices: advertising,[226] media, activist movements, influencer marketing, education, psychology (e.g. CBT), law, sociology, etc. The outsourcing choice may, however, depend on different factors and may change over time and across different contexts. The particular implementation choice may also be subjected to careful scrutiny and informed by empirical data. For instance, as part of an engineering project, the media might explicitly advocate that a particular generic is being asserted, rejected, or refrained from being asserted/rejected. This could involve the giving of reasons to try to persuade the public even though sometimes it could be done implicitly or by deferring to the experts or background theories. It could also be done in synergy with certain e.g. social movements. There is no one size fits all strategy here. At the end of the day, what is important in the present context is for the audience to form or not to form certain generic judgments which may lead to pernicious, beneficial or neutralizing inferences.

6.5.6 No Unitary Solution

My proposal does not offer a *unitary solution* when it comes to particular instances of pernicious generics. It is worth reminding the reader that the main goal of engineering projects is to engineer (pernicious/beneficial/neutralizing) *effects* of generic judgments, and not for everyone to share the same propositional attitudes or judgments about generics. With this in mind, it is important to stress that different instances of generics may require different particular solutions depending on various

---

[225] In "a paradigmatic illocutionary event, a speaker has a choice of which if any speech act to perform and her addressee will do her charitable best to discern that speaker's intentions and, where necessary, which conventions she may be invoking" (Green 2020).

[226] Compare Nado's suggestion for implementing the change in classificatory procedures: "However, ultimately this is all to do with 'advertising'. One nice feature of the move away from a semantic approach is that advertising can be viewed as secondary to engineering proper—that is, secondary to the actual discovery or creation of a classification procedure that effectively fills a needed practical role" (Nado 2020, 24–25).

factors such as: context,[227] time, resources, the scope of the implementation, contents of a particular generic proposition, audience's background believes, audience's psychological profile, the primacy of target effects of generics, the role a particular solution may play in other engineerings, etc. For example, paying attention to context sensitivity when engineering generic judgments, may affect engaging in local vs. global[228] engineering. Furthermore, *assertion* of a certain generic could bear different level of contextual injustice[229] which could, in result, have different impact or require different ameliorative strategies, depending on the context. This could, for instance, result in *refraining from asserting* (e.g. Blacks are violent) in some cases and *refusing* (e.g. "Women are submissive") or *asserting* (e.g.'Girls are good at math') in some other cases. Moreover, refraining from judgment, e.g. "Muslims are terrorists," or "Blacks are violent" could be an intermediary state in cases where speakers are highly radicalized. Lastly, there might also be a difference in strategic planning and implementation when applied to children as opposed to adults given the difference in their e.g. psychological makeup, background assumptions, knowledge, trust, etc.[230]

6.6 Re-evaluation

Re-evaluation[231] can serve several roles such as to: (i) identify and describe the post-implemented change in propositional attitudes towards generics (e.g. from *asserting* a generic proposition "Muslims are terrorists" to *rejecting* it); (ii) account for potential changes in evaluation approach, e.g. in case when there has been a change in what counts as pernicious, beneficial or neutralizing effect for a particular generic proposition; (iii) detect potential errors during the strategic planning stage, e.g. in case if a better strategy for a particular generic proposition has been developed or overlooked; (iv) detect potential errors during the implementation stage, e.g. in case when a particular strategy is not implemented correctly (in a certain context); (v) test the feasibility of the proposed strategy and the feasibility of its implementation. Finally, this stage could be seen as the beginning of a new recursive cycle.

---

[227] "Even though, psychological essentialization may be a universal human phenomenon, there is significant cross-cultural variation as to which kinds are seen as essentialized" (Leslie 2017, 415).

[228] E.g. engineering judgments of all speakers, or engineering judgments of particular (group of) speakers. For *global* vs. *local* engineering more generally, see Chapter 2, Section 2.3.2.4.

[229] See Ichikawa (2020) for contextual injustice.

[230] E.g. here is how (Pollock 2020) spells out the role of trust in forming believes and concepts: "if someone in whom you have sufficient trust tells you that an octopus is actually a fungus, there is not much you can do to prevent your conceptual network being updated. However, this is simply for the reason that we cannot choose what we believe" (Pollock 2020, 12). One may argue that trust could play a similar role in the case of changing propositional attitudes.

[231] For more about *re-evaluation* stage more generally see Chapter 2, Section 2.3.2.5.

6.7 Benefits

In this section, I consider three key benefits which make the Speech Act Approach a plausible alternative approach to achieve *engineering generic judgments*.

6.7.1 The Speech Act Approach Does Not Depend on Any Particular View About the Semantics, Epistemology or Metaphysics of Generics

Speakers often *assert* or *refuse* generics, *independently* of whether they are in a position to *know* the truth value of a generic proposition in question. In other words, when speakers *assert* or *refuse* a certain generic proposition, this does not automatically mean that they *know* whether a particular statement is *true* or *false*, nor that a particular statement *is*, in fact, *true* or *false*. For example, speakers can *assert* or *refuse* a generic such as "Muslims are terrorists." However, mere *assertion* or *refusal* does not mean that the proposition in question is *true* or *false*. *Speakers' assertion* or *refusal of a certain generic* is not a reliable indicator[232] of (their knowledge of) the truth value of generics since semantics, epistemology and pragmatics of generics can come apart (see Haslanger 2011; Leslie 2017; Chapter 5, Section 5.3; Fuš manuscript b).

In this chapter, I argued that while the effects stemming from our judgments about generic proposition may be independent of their semantics (truth value), epistemology (speaker's knowledge of the truth value) and metaphysics (what generic properties, if any, a generic proposition picks out), they are sensitive to a change in speech acts (such as assertion, refusal, or refraining from assertion/refusal) towards generic propositions.[233] The Speech Act Approach that I proposed aims at achieving the change in generic judgments independently of changing the current truth value,[234] audience's knowledge about current truth value, and generic properties (if any) that the generic proposition picks out. This makes engineering generic judgments possible even if the semantic value were unknowable, or even if semantic, epistemic or metaphysical change were hard or impossible. Thus, the good news is that even if engineering generic judgments via semantic, epistemic or metaphysical engineering may not be

---

[232] It is often taken for granted that competent speakers of a language are sufficiently good at identifying propositions that are being asserted and beliefs that are being expressed by utterances. Given that a semantic theory for a certain language belongs to a larger theory that interprets the assertions and beliefs of its speakers, any given semantic theory can be empirically tested. The confirmation or disconfirmation of a particular semantic theory depends on the extent to which assignments of semantic content given by a semantic theory lead to verifiably correct or incorrect characterizations of speakers' assertions and beliefs. Given the above, while theorizing about the truth-conditions of generics the semanticists about generics should utilize *speakers' intuitions* about the truth or falsity of generics by looking at speakers' *propositional attitudes* towards generics. However, Leslie (2017) goes against speakers' intuitions in order to offer truth-conditions for striking property generics. Recall, Leslie (2017) argues that certain striking property generics that speakers *accept* as *true* are *actually false*, e.g. pernicious generics such as "Muslims are terrorists." Her explanation of this 'abnormality' is that they *accept* such generics as true because they *psychologically essentialize* the kind in question (e.g. Muslims) as having the predicated property (e.g. being a terrorist).

[233] Relatedly, I argued that it is irrelevant whether generic judgments are generated through false or true predication, false conversational or relevance implicatures, or psychological essentializing. Instead they may only be facilitated through some of the above mentioned processes.

[234] It is worth noticing that my proposal is more permissive than Langton, Haslanger, and Anderson's (2012, 765) who call for a rejection of asserted racial generics that apart from being "politically problematic" are also *false*.

fruitful, it could be fruitful via pragmatic engineering. Moreover, given that the Speech Act Approach does not rely on any semantic, epistemic or metaphysical properties of generics, it also eschews the already complex talk about the semantics, epistemology and metaphysics of generics.

6.7.2 The Speech Act Approach Is More Feasible to Implement Than the Proposed Alternatives

In Section 6.4.1, I characterized Leslie's approach that results in abolishing generics about social groups as self-undermining when it comes to its implementation given her hypothesis that generics are a fundamental way of generalizing. Relatedly, projects which, on the one hand, incorporate the view that pernicious generics go hand in hand with false predication whereas, on the other hand, also rely on the *truth assertion norm* when it comes to their amelioration (see Ritchie 2019), could sometimes turn out to be *counterproductive* for getting rid of pernicious generic judgments. For example, it might turn out that, after all, "Blacks are violent" or "Women are submissive" is *true* in which case, following the truth assertion norm, one should assert these generics. In such cases, a defender of this approach might want to retain her commitments by trying to change the truth value of the pernicious generic proposition that is true. However, semantic change of generics would, arguably, require more resources than changing the speech acts towards the same generic. The Speech Act Approach avoids this problem altogether since it is not governed by truth assertion norm.[235]

Furthermore, proposals such as Haslanger's and Saul's which rely heavily on an active rational reflection on the part of the speakers are arguably harder to implement in comparison to the Speech Act Approach. The Speech Act Approach does not require from neither the implementator nor the audience to know whether a particular generic proposition is true or false. More importantly, it does not require from the audience to have justification for why they have particular propositional attitudes towards generic propositions. The Speech Act Approach rests primarily on the expertise of engineers and the authority of implementators, instead on the cognitive and rational capacity of the audience to critically asses which generic judgments may be most apt for changing the effects of generic judgments. This feature makes the Speech Act Approach especially feasible among the children. Even though it may be independently beneficial for the children (likewise for the adults) to develop the skills of critical thinking in a similar way Haslanger (2011) and Saul (2017) propose, children might sometimes lack the higher order reasoning or might find it cognitively too taxing to engage in rational reflection (see Section 6.4.3). Moreover, the Speech Act Approach should be easier to implement among the children since the empirical research shows that children are especially susceptible to learning new generics and

[235] This said, a possible change in the truth value of generics such as "Blacks are violent" as a result of e.g. changing the underlying conditions of oppression is compatible with the Speech Act Approach.

adopting propositional attitudes towards generics when they are taught.[236] In result, this could block certain perniciousness even before it was created or transmitted. This said, even though the Speech Act Approach does not require the strategies proposed by Haslanger (2011) and Saul (2017), it can sometimes be facilitated by applying the strategies like this that require rational reflection on behalf of the audience in order to achieve the change in propositional attitudes towards generics.

6.7.3 The Speech Act Approach Provides a Better Framework for Utilizing Generics in Order to Achieve Beneficial or Neutralizing Effects

If Leslie's (2017) is right, then we should avoid all generics about social groups.[237] According to her, we should avoid not only generics such as "Blacks are violent," "Muslims are terrorists," "Women are submissive" but also generics such as "Girls are good at math," "Girls are tough,""Girls should be allowed to play football," "Boys shouldn't be picked on for liking pink," "Boys cry too," "Boys like pink, too," "Girls play football," "Girls can do anything," "White silence is violence," "Black lives matter," "Blue lives matter."

However, as already noted by Saul (2017, 12–13): "Although Leslie and Haslanger's focus is on the ways that generics can serve to perpetuate racism, sexism and other ills, it also seems to me worth noting the ways that they can serve to fight these ills." At least some of the above generics can, arguably, have beneficial or neutralizing effects, especially when thought and communicated to children.[238] They can help us fight pernicious generic judgments which may help to combat certain stereotypes and prejudices generic judgments may lead to. For example, generics such as "Women are expected to want children," "Black people face discrimination," "Gay people are subjected to violence," "Muslims are profiled by airport security" could be utilized by campaigners for social justice, as part of a social critique. In certain contexts, generics can also be used to stress or raise awareness about something that needs to be changed, e.g. "Boys will be boys!" or "Women all over America are fighting greasy pots and pans." They can also be used to express display rules such as "People with cancer should be

[236] "Children and adults who heard the Zarpies described with generic noun phrases were significantly more likely to essentialize Zarpies. Moreover, in a follow-up study, we found that when parents themselves were led to essentialize Zarpies, they produced significantly more generics about Zarpies when discussing the group with their children. Taken together, our studies suggest that generics may be a means by which social essentialist beliefs are transmitted across generations, as parents who themselves hold more essentialist beliefs about a group produce more generics, which in turn leads young children to form essentialist beliefs of their own." (Leslie 2017, 416)

[237] Apart from Leslie 2017, the proposals of e.g. Haslanger 2011; Langton, Haslanger, and Anderson 2012; Wodak, Leslie, and Rhodes 2015; Wodak and Leslie 2017 can also be seen as resulting in avoidance of racial and gender generics (see Ritchie 2019).

[238] As Saul (2017) points out: "according to Leslie, they make use of a form of generalisation that comes especially easily and naturally to children. (…) Depriving us of [these] statements (…) is depriving us of some very important weapons in our anti-prejudice arsenal" (Saul 2017, 13).

optimistic."[239]

Similarly, Ritchie (2019) argues against a general prohibition proposed by Leslie (2017) since generics such as "Blacks face economic, legal, and social discrimination" or "Women are expected to want children" when "describing forms of structural oppression generic generalizations can be the most accurate and effective linguistic tools to describe social reality. There are some racial and gender generics that we do not need to avoid and that perhaps we even ought to use" (Ritchie 2019, 39).[240]

Pace Leslie (2017), I agree with both Saul (2017) and Ritchie (2019) that we should not avoid nor reject all generics about social groups. The Speech Act Approach provides a framework that is compatible not only with the amelioration of pernicious generics but also with the endorsement of generics that could help us achieve the beneficial and neutralizing effects, regardless of their truth value and the audience's knowledge of their truth value. Furthermore, what makes the Speech Act Approach a better competitor is that, as opposed to Saul's (2017) approach, in order to achieve the beneficial and neutralizing effects the Speech Act Approach does not necessarily require the speaker's rational reflection. As opposed to Ritchie's (2019) approach, the Speech Act Approach allows for the beneficial and neutralizing effects to be achieved by utilizing generics regardless of their truth value and descriptive accuracy.

6.8 Ethical/Epistemic Dilemma

6.8.1 Epistemic Function

The Speech Act Approach aims at changing audience's assertion, refusal, or refrainment from assertion/refusal of a generic proposition regardless of whether what is being asserted, refused or refraining from being asserted/refused is true or false. In principle, the Speech Act Approach, allows for

---

[239] However, introducing beneficial generics comes with its own perils. Below are a couple of examples.

Sometimes a particular (historical) context and what has been implied may also play a significant role in these effects. For instance, "Boys love pink too" can be seen as beneficial for those who endorse a stereotype that "Boys don't love pink" but potentially harmful for those who hear that generic for the first time since it may imply that boys don't love pink or, perhaps even worse, that only girls love pink.

Some generics could be pernicious with respect to one aspect but beneficial with respect to another. For example, "Asians are smart" could be perceived as racist but also can be perceived as fostering a positive stereotype.

Sometimes endorsing certain generics could have the beneficial effects on one subgroup, while having unwanted effects on the other. For instance, endorsing "Girls are good at math" may affect some of those who identify as girls and believe that they are not good at math to infer that they are less of a 'proper girl.' (Thanks to Mark Pinder for pointing this out). Interestingly enough, generic "Girls are good at math" has been proposed to achieve quite the opposite, i.e. to mitigate the *stereotype threat* taken to be originating from the prejudice that girls are bad at math since it has been shown that those who self-ascribe this generic tested worse in math than those who didn't.

In such (but perhaps all) cases, I would propose that the engineer should engage in benefit-cost analysis. According to my view, the decision should also be informed by psychology and empirical data (if possible) before coming to a conclusion which formulations would be better, all things considered. In addition, one could also introduce a parallel generic such as e.g. "People are not obliged to be good at math" to target the above mentioned subgroup for which the new generic might have pernicious effects.

[240] Ritchie (2019) criticizes Saul's (2017) view on the usefulness of generics in social justice work as incomplete because Saul (2017) does not offer "an explanation for the efficacy of generics and does not claim that they are *true* or *accurate*. If social generics can be socially and politically efficacious, what explains their effectiveness?" (Ritchie 2019, 37; italics mine)

a change of an assertion "Blacks are violent" to refusal of the same generic even if one does not know whether the asserted generic is actually true of false. As such, The Speech Act Approach may be open to the following cluster of objections: *misrepresentation, asserting without knowing, lying, imposing false beliefs, being irrational*.

A direct motivation for these worries comes from the long standing idea about the importance of epistemic function of our representational devices. In the context of engineering in philosophy, one can roughly distinguish between the following three three positions: *conservative* (see Simion 2018), *moderate* (see Scharp 2020), and *liberal* (see Nado 2019).

A version of the *conservative* view, defended by Simion (2018) states: "Concepts, just like beliefs, are representational devices, their function is an epistemic one: to represent the world. In virtue of this function, concepts will be properly functioning when responsive to epistemic reasons, and malfunctional when responsive to practical reasons. Concepts will be good concepts qua concepts when they are epistemically good" (Simion 2018, 923).

A version of the *moderate* view is endorsed by Scharp (2020) who claims: "I tend to err on the side of letting a thousand flowers bloom rather than figuring out from the armchair which conceptual engineering projects are kosher. Still, I can envision balking at certain proposals if they were to, for example, promote conceptual confusion or inconsistency for political gain. I think conceptual engineering should always make our conceptual scheme better for us and the concepts we use better for what we use them for. I like to think of conceptual engineering as a wide category, but it certainly has limits" (Scharp 2020, 410).

A version of a *liberal* response, which is compatible with what Isaac (manuscript) calls a *psychological take* on our representational devices takes it that: "… while concepts enable us, cognitive agents, to execute the cognitive task we use them for, they are not primarily driven by truth requirements and further constrained by epistemic standards—but instead only by functional, cognitive efficacy (see Nado 2019)—whereas, on the philosophical take, while aiming at delivering truths about the world, they ultimately always are" (Isaac manuscript).

When it comes to engineering generic judgments, I take a rather *liberal* stance. In the rest of this section, after putting the above cluster of worries in the context of a larger literature on the *ethical/epistemic dilemma*, I offer some additional support for why, in the case of *engineering generic judgments* particularly, one is (more) justified to sidestep the ethical/epistemic dilemma.

6.8.2 A Dilemma

In the background of the above mentioned cluster of worries lies the *ethical/epistemic dilemma.* The dilemma boils down to a conflict between one's ethical vs. epistemic aims. This type of dilemma can be

traced at least back to Pascal. As Scharp (2020, 409) puts it: "Pascal's reasons for belief in God are entirely focused on the consequences of having that belief, rather than on whether the belief is true." Modern ethical/epistemic dilemmas can be found in the literature on e.g. *implicit bias* (see Egan 2011; Gendler 2011) and *the value of positive illusions*[241] (see Taylor and Brown 1988). In particular, the objector may claim that *implicit bias* or *positive illusions* create a conflict between our ethical and epistemic aims. For example, in the case of implicit bias, when faced with the fact that one cannot escape stereotyping, one may consider whether to give up social stereotypes on ethical grounds, and, thus, open a dilemma between our knowledge and our values. As Gendler (2011) stresses:

> … if you live in a society structured by racial categories that you disavow, either you must pay the epistemic cost of failing to encode certain sorts of base-rate or background information about cultural categories, or you must expend epistemic energy regulating the inevitable associations to which that information—encoded in ways to guarantee availability—gives rise. (Gendler 2011, 37) [242]

One solution to the above ethical/epistemic dilemma, endorsed by Egan (2011), is to accept that ethical values are worth epistemic sacrifices. Another solution, proposed by Madva (2016), is the "situation-specific regulation of stereotype accessibility," based on agents' goals and habits (see Moskowitz and Li 2011), relying on an argument that "it is not social knowledge *per se* that has costs, but rather that the accessibility of social knowledge in the wrong circumstances has cognitive costs" (Brownstein 2019).

In the context of conceptual engineering, Scharp (2020) revives this dilemma in a somewhat different setting. He contrasts two kinds of reasons for changing one's representational devices: i.e. *reasons of the wrong kind* and *reasons of the right kind*. First, he claims that: "Reasons for belief of the right kind are those relevant to the truth of the belief, whereas reasons for having a belief, which are irrelevant to the truth of the belief, are reasons of the wrong kind" (Scharp 2020, 409). Second, he offers Pascal's wager as an example of the reasons of the wrong kind and compares it to conceptual engineering examples claiming that:

> … the reason for changing our concept of time or our concept of rights is not a reason to think the new concept is a good one. Instead, the reason given is maintaining certain moral and social commitments (e.g., pro-choice and anti-infanticide). This is a reason to change the concepts I use, but it is not a reason to think that the new ones are valuable or right for the job. (Scharp 2020, 409)

Third, he contrasts the above examples which involve cases of *beliefs* with conceptual engineering

---

[241] E.g. beliefs like "I am brilliant!" which may promote well-being despite being false (at least in some cases).

[242] Gendler (2011) offers the following four examples: cross-race recognition deficits; stereotype threat; cognitive depletion following interracial interactions; and "forbidden" base-rates.

that is an *activity*, and stresses that, in the case of actions, there may not be reasons of the wrong kind. As he points out:

> … there is no difference in kind between volunteering at a homeless shelter because I want to help the homeless and volunteering at a homeless shelter due to receiving some incentive. Hence, it might be that there are no reasons of the wrong kind to promote moral rightness. So promoting moral rightness is on par with any other reason when it comes to adopting a certain concept of time or a certain concept of persistence, or a certain concept of rights. (Scharp 2020, 410)

Finally, Scharp takes a moderate view by concluding that his attitude on the matter has not been "entirely stable" and that it may depend on the case in question (see Scharp 2020, 410).

6.8.3 Sidestepping the Ethical/Epistemic Dilemma for Generics

In addition to the above mentioned general responses to the ethical/epistemic dilemma, I offer three additional reasons for sidestepping the ethical/epistemic dilemma in the case of generics in particular.

6.8.3.1 We Might Not Be in a Position to Know the Truth Value of a Generic Proposition

The side-stepping of the dilemma rests on the idea that when the epistemic role is out of reach, we might be more justified to focus on the ethical role of our representational devices. Some possible support that this may be so in the case of generics are arguments that semantics, epistemology and pragmatics of generics can come apart and that speakers' intuitions should not be trusted since we are often bad at thinking about truth-conditions of generics (see Haslanger 2011; Leslie 2017); incapacity of the state of the arts to give a unifying semantic theory of generics (see Chapter 5); possible antirealism about the generic truth (see Fuš manuscript a,b; Chapter 5); Arbitrary Reference (an epistemicist view of generics advocated in Fuš manuscript b). Additionally, one may argue that there is an analogy between generics and *underspecified*, *vague*, and *indeterminate* representational devices. Given that the latter have often been utilized because of their positive effects, one may, analogously, argue that we could also utilize generics for similar purposes too.

6.8.3.2 Non-epistemic Goals Such as Moral, Political, or Social Can at Least Sometimes Trump the Epistemic Goals

In certain cases, when there is a *possibility* rather than *certainty* of an epistemic cost, and when what leads to it is done *unknowingly* rather than *intentionally*, one may argue that non-epistemic positive goals such as moral, political, or social ones can trump the epistemic ones. Given the above mentioned possibility of deep ignorance about the truth value of generics, one may argue that *asserting* certain false

generics when done *unknowingly* and *unintentionally* is permissible. Furthermore, given that the possible ignorance about the truth value of generics puts one in an epistemic position where there is no intentional departure from the truth of a generic, engineering generic judgments may be seen as less dishonest or less ethically questionable than phenomena such as *fake news, deceiving, manipulating, bullshit, post-truth* or *lying*.

6.8.3.3 The Irrationality Worry Does Not Apply to Generics in the Same Way It Does to the Universal and Existential Statements

The *irrationality worry*, which rests on the truth assertion norm, says that it is irrational to assert falsehoods and it is rational to assert truth.[243] One component of a rationality test (a strategy often used in the case of *universally* and *existentially* quantified statements) is to challenge speakers' believes in the presence of counterexamples. When applied to *universally* and *existentially quantified statements,* this strategy is rather successful since the speakers can have clear knowledge of what it takes for the universally and existentially quantified statements to be true. For instance, imagine this strategy being applied to a universally quantified sentence "All ravens are black" (that was previously *accepted as true* before it was discovered and pointed out to the speakers that there are albino ravens). On the other hand, generics are persistent in the face of counterexamples.[244] One result of this is that they are more difficult to challenge.[245] Another, more important one in the present context, is that we do not have a good enough rationality test for generics. Thus, if speakers' *initial tendencies* to *assert* or *refuse* certain generics are insensitive to the rationality test that complies with the truth assertion norm, then the same should apply to their *engineered tendencies*, namely we cannot expect that their engineered tendencies should be sensitive to the rationalist test that complies with the truth assertion norm in order for the speakers to be considered rational.

6.9 Conclusion

In this chapter, I worked towards achieving two goals. First, I identified and critically examined current approaches to pernicious generics (Haslanger 2011, Leslie 2017, Saul 2017, Ritchie 2019) as

---

[243] One can find a way around this worry by agreeing with Scharp (2020) who argues that even defective concepts may still be used by rational agents because they can be useful: "The analogy I like to use is the concept of mass in classical mechanics. Mass is inconsistent concept but is still extraordinarily useful to use it for all kinds of things from building houses to landing robots on comets" (Scharp 2020, 412).

[244] "accepting a generic disposes people—all else being equal—to believe that an arbitrary instance of the kind will have the property, as has long been suggested in the defeasible reasoning literature (e.g., Pelletier and Asher 1997)." (Leslie and Lerner 2016)

[245] Langton, Haslanger and Anderson's (2012, 764) claim that striking property generics are 'easy to accept, hard to refute.' They see the reason for this effect in psychological slippage between *striking property* and *characteristic* generics. As they put it: "Psychological research shows, moreover, that *striking property generics* are *quickly accepted based on few instances*, whereas *characteristic generics*, once accepted, are *resistant to refutation*. Given the slippage between these two forms, properties not indicative of a general pattern get attributed to a group and then stick (Cimpian, Brandone and Gelman, [2010]). *Easy to accept, hard to refute* is not an epistemically or politically promising combination" (Langton, Haslanger and Anderson 2012, 764; italics mine).

domain-specific accounts of philosophical engineering of generics. Second, I developed the Speech Act Approach as an alternative domain-specific account to philosophical engineering of generics.

This chapter makes a broader point that a lot of our interest in ameliorating the effects of generics has been misplaced since debates in semantics, epistemology and metaphysics of generics may not offer us necessary nor sufficient tools to change the effects of generics. I argued that *false predication* (see Leslie 2017), *true predication* (see Ritchie 2019), *false conversational implication* (see Haslanger 2011), *false relevance implication* (see Saul 2017), or *psychological essentializing* are neither sufficient nor necessary for the effects of generics. Instead, I argued, they stem, most fundamentally, from our *generic judgments* which may, nevertheless, be facilitated through some of the above mentioned processes. Given the close connection between generic judgments and speakers' propositional attitudes towards generic propositions, I suggested *engineering generic judgments* that operates on utilizing speech acts in order to change the propositional attitudes towards generic propositions.

My proposal, however, builds on some of the key insights from the debate about engineering pernicious effects of generics. In particular, by acknowledging the possibility that the semantics, epistemology and pragmatics of generics can come apart (see Haslanger 2011) and by acknowledging the phenomenon of psychological essentializing (see Leslie 2017), I argued that the engineers (rather than the speakers (see Saul 2017)), should use rational reflection to come up with particular strategies for changing the pernicious and bringing about beneficial and neutralizing effects of generics (see Saul 2017; Ritchie 2019).

BIBLIOGRAPHY

Alcoff, L. 2000. Is Latino/a Identity a Racial Identity? In *Hispanics and Latinos in the United States: Ethnicity, Race, and Rights,* eds. Gracia, J J. E. and P. De Grieff. New York: Routledge.

Andow, J. 2020. Fully Experimental Conceptual Engineering. *Inquiry,* 10.1080/0020174X.2020.1850339.

Andreason, R. 2000. Race: Biological Reality or Social Construct? *Philosophy of Science*, 61, Supplementary Volume: S653–666.

Appiah, K. A. 1996. Race, Culture, Identity: Misunderstood Connections. In *The Tanner Lectures on Human Values XVII*, ed. G.B. Peterson. Salt Lake City: University of Utah Press.

Asher, N. and F.J. Pelletier. 2012. More Truths About Generic Truth. In *Genericity*, eds. Mari, A., Beyssade, C. and F. D. Prete. Oxford: Oxford University Press.

Bach, K. 1994. Conversational Impliciture. *Mind and Language,* 9(2): 124–62.

—— 2000. Quantification, Qualification, and Context: A Reply to Stanley and Szábo. *Mind and Language,* 15(2&3): 262–83.

Baldwin, D.A., Markman, E. and R. Melartin. 1993. Infants' Ability to Draw Inferences About Nonobvious Object Properties: Evidence from Exploratory Play. *Child Development*, 64(3): 711–28.

Ball, D. 2020. Revisionary Analysis without Meaning Change (Or, Could Women Be Analytically Oppressed?). In *Conceptual Engineering and Conceptual Ethics*, eds. Burgess A., Cappelen H. and D. Plunkett. Oxford: Oxford University Press.

Barkow, J.H., Cosmides, L. and J. Tooby. 1992. *The Adapted Mind: Evolutionary Psychology and the Generation of Culture*. New York: Oxford University Press.

Barrett, J.L. 2007. Is the Spell Really Broken? Bio-psychological Explanations of Religion and Theistic Belief. *Theology and Science,* 5(1): 57–72.

Belleri, D. 2018. Two Species of Merely Verbal Disputes. *Metaphilosophy,* 49(5): 691–710.

—— Manuscript. Deflationary Projects in Ontology and Conceptual Engineering.

Beaney, M. 2018. Analysis. In *The Stanford Encyclopedia of Philosophy* (Summer Edition), ed. E.N. Zalta. URL = <https://plato.stanford.edu/archives/sum2018/entries/analysis/>.

Berners-Lee, T. 2003. Re: New Issue—Meaning of URIs in RDF Documents. *Message to www-tag@w3.org Mailing List* [Internet]. Available from: http://lists.w3.org/Archives/ Public/www-tag/ 2003Jul/0158.html [Accessed 11 January 2021].

Blackburn, S. 1999. *Think: A Compelling Introduction to Philosophy.* Oxford: Oxford University Press.

Block, N. 1995. On a Confusion About a Function of Consciousness. *Brain and Behavioral Sciences,* 18: 227–247.

Brandom, R. 1994. *Making It Explicit.* Cambridge, MA: Harvard University Press.

—— 2001. Modality, Normativity, and Intentionality. *Philosophy and Phenomenological Research,* 63(3): 587–609.

Brandone, A., Gelman, S.A. and J. Hedglen 2015. Children's Developing Intuitions About the Truth Conditions and Implications of Novel Generics Versus Quantified Statements. *Cognitive Science*, 39(4): 711–738.

Breckenridge, W. and O. Magidor. 2012. Arbitrary Reference. *Philosophical Studies*, 158(3): 377–400.

Brigandt, I. 2010. The Epistemic Goal of a Concept: Accounting for the Rationality of Semantic Change and Variation. *Synthese,* 177(1): 19–40.

Brigandt, I. and E. Rosario. 2020. Strategic Conceptual Engineering for Epistemic and Social Aims. In *Conceptual Engineering and Conceptual Ethics*, eds. Burgess A., Cappelen H. and D. Plunkett. Oxford: Oxford University Press.

Brownstein, M. Implicit Bias. 2019. In *The Stanford Encyclopedia of Philosophy* (Fall Edition), ed. E.N. Zalta. URL = <https://plato.stanford.edu/archives/fall2019/entries/implicit-bias/>.

Brun, G. 2016. Explication As a Method of Conceptual Re-engineering. *Erkenntnis,* 81(6): 1211– 1241.

—— 2017. Conceptual Re-engineering: From explication to reflective equilibrium. *Synthese*, 197: 925–954.

Burgess, J. 2014. Review of Scharp (2013). *Studia Logica,* 102: 1087–1089.

Burgess, A., Cappelen, H. and D. Plunkett. 2020. *Conceptual Engineering and Conceptual Ethics.* Oxford: Oxford University Press.

Burgess, A. and D. Plunkett. 2013a. Conceptual Ethics I. *Philosophy Compass,* 8(12): 1091– 1101.

—— 2013b. Conceptual Ethics II. *Philosophy Compass,* 8(12): 1102– 1110.

—— 2020. On the Relation Between Conceptual Engineering and Conceptual Ethics. *Ratio*, 1–14.

Buss, D. M. 1999. *Evolutionary Psychology: The New Science of Mind*. Boston: Allyn and Bacon.

Cappelen, H. 2017. Why Philosophers Shouldn't Do Semantics. *Review of Philosophy and Psychology*, 8(4): 743–762.

—— 2018. *Fixing Language: An Essay on Conceptual Engineering*. Oxford: Oxford University Press.

—— 2020. Conceptual Engineering: The Master Argument. In *Conceptual Engineering and Conceptual Ethics*, eds. Burgess A., Cappelen H. and D. Plunkett. Oxford: Oxford University Press.

Cappelen, H. and E. Lepore. 2001. Insensitive Quantifiers. In *Meaning and Truth: Investigations in Philosophical Semantics*, eds. Campbell, J.K., O'Rourke, M. and D. Shier. New York, NY: Seven Bridges Press.

—— 2002. Indexicality, Binding, Anaphora, and A Priori Truth. *Analysis*, 62(4): 271–81.

—— 2005. *Insensitive Semantics*. Oxford: Blackwell.

Cappelen, H. and D. Plunkett. 2020. Introduction: A Guided Tour of Conceptual Engineering and Conceptual Ethics. In *Conceptual Engineering and Conceptual Ethics*, eds. Burgess A., Cappelen H. and D. Plunkett. Oxford: Oxford University Press.

Carlson, G. 1977. *Reference to Kinds in English*. Ph.D. dissertation. University of Massachusetts, Amherst.

—— 1989. On the Semantic Composition of English Generic Sentences. In *Properties, Types and Meanings, Vol. II: Semantic Issues*, eds. Chierchia, G., Partee, B.H. and R. Turner. Dordrecht: Kluwer.

—— 2011. Genericity. In *Semantics: An International Handbook vol. II*, eds. Maienborn, C., von Heusinger, K. and P. Portner. Berlin: de Gruyter: Berlin.

Carnap, R. 1947. *Meaning and Necessity. A Study in Semantics and Modal Logic.* The University of Chicago Press.

—— 1950. *The Logical Foundations of Probability*. The University of Chicago Press, Chicago Illinois.

—— 1963. Replies and Systematic Expositions. In *The Philosophy of Rudolf Carnap*, ed. P.A. Schilpp. La Salle: Open Court.

Carston, R. 1996. Metalinguistic Negation and Echoic Use. *Journal of Linguistics*, 25: 309–30.

Cavedon, L. and S. Glasbey. 1994. Outline of an Information-Flow Model of Generics. *Acta Linguistica Hungarica*, 42.

Chalmers, D. 2011. Verbal Disputes. *Philosophical Review*, 120(4): 515–566.

—— 2020. What is Conceptual Engineering and What Should It Be? *Inquiry*, doi:10.1080/0020174X.2020.1817141.

Chihara, C. 1979. The Semantic Paradoxes: A Diagnostic Investigation. *Philosophical Review,* 88: 590– 618.

Cimpian, A., Brandone, A.C. and S.A. Gelman. 2010. Generic Statements Require Little Evidence for Acceptance but Have Powerful Implications. *Cognitive Science*, 34(8): 1452–1482.

Cimpian, A. and L.C. Erickson. 2012. Remembering Kinds: New Evidence that Categories Are Privileged in Children's Thinking. *Cognitive Psychology*, 64(3): 161–185.

Cimpian, A., Gelman, S.A. and A.C. Brandone. 2010. Theory-based Considerations Influence the Interpretation of Generic Sentences. *Language and Cognitive Processes*, 25(2): 261–276.

Clark, A. and D. Chalmers. 1998. The Extended Mind. *Analysis*, 58(1): 7–19.

Cohen, A. 1996. Think Generic: The Meaning and Use of Generic Sentences. Ph.D. dissertation, Carnegie Mellon University.

—— 1999. Generics, Frequency Adverbs and Probability. *Linguistics and Philosophy*, 22(3): 221–253.

—— 2004. Generics and Mental Representation. *Linguistics and Philosophy*, 27(5): 529–556.

Cohen, G. A. 2000. *If You're an Egalitarian, How Come You're so Rich?* Cambridge, MA: Harvard University Press.

Craig, E. 1990. *Knowledge and the State of Nature: An Essay in Conceptual Synthesis*. Oxford: Clarendon.

Creath, R. 1990. *Dear Carnap, Dear Van: The Quine-Carnap Correspondence and Related Work.* University of California Press.

Csibra G. and R. Shamsudheen. 2015. Nonverbal Generics: Human Infants Interpret Objects as Symbols of Object Kinds. *Annual Review of Psychology,* 66(1): 689–710.

Davidson, D. 1970. Mental Events. In *Experience and Theory*, eds. Foster, L. and J. W. Swanson. Clarendon Press.

—— 1973. Radical Interpretation. *Dialectica*, 27(3&4): 313–328.

—— 1974. Belief and the Basis of Meaning. *Synthese*, 27(3&4): 309–323.

—— 1978. What Metaphors Mean. *Critical Inquiry,* 5(1): 31–47.

Declerk, R. 1986. The Manifold Interpretations of Generic Sentences. *Lingua,* 68: 149–188.

Deutsch, M. 2020a. Speaker's Reference, Stipulation, and a Dilemma for Conceptual Engineers. *Philosophical Studies,* doi: 177: 3935–3957.

—— 2020b. Trivializing Conceptual Engineering. *Inquiry*, doi: 10.1080/0020174X.2020.1853343.

Devitt, M. 1996. *Coming to Our Senses.* Cambridge: Cambridge University Press.

Dorr, C. and J. Hawthorne. 2014. Semantic Plasticity and Speech Reports. *Philosophical Review*, 123(3): 281–338.

Dummett, M. 1978. *Truth and Other Enigmas*. London, Duckworth.

Dutilh Novaes, C. 2016. Conceptual Genealogy for Analytic Philosophy. In *Beyond the Analytic-Continental Divide: Pluralist Philosophy in the Twenty-First Century*, eds. Bell, J.A., Cutrofello, A. and P.M. Livingston. New York: Routledge.

—— 2018. Carnapian Explication and Ameliorative Analysis: A Systematic Comparison. *Synthese*, 1–24.

Dutilh Novaes, C. and E. Reck. 2017. Carnapian Explication, Formalisms as Cognitive Tools, and the Paradox of Adequate Formalization. *Synthese,* 194(1): 195–215.

Egan, A. 2011. Comments on Gendler's 'The epistemic costs of implicit bias'. *Philosophical Studies*, 156: 65–79.

Eklund, M. 2002. Inconsistent Languages. *Philosophy and Phenomenological Research,* 64(2): 251–275.

—— 2014. Replacing Truth? In *Metasemantics: New Essays on the Foundations of Meaning,* eds. Burgess, A. and B. Sherman. Oxford: Oxford University Press.

—— 2015. Intuitions, Conceptual Engineering, and Conceptual Fixed Points. In *The Palgrave Handbook of Philosophical Methods*, ed. C. Chris. Palgrave Macmillan Houndsmills.

Fassio, D., and R. McKenna. 2015. Revisionary Epistemology. *Inquiry*, 58(7–8): 755–779.

Fisher, J. C. 2015. Pragmatic Experimental Philosophy. *Philosophical Psychology,* 28(3): 412–433.

Fodor, J. 2008. *LOT 2: The Language of Thought Revisited*. Oxford: Oxford University Press.

Frege, G. 1892. On Sense and Reference. Translated in *Philosophical Review*, 57(3)(1948), 209–30.

Fricker, M. 2007. *Epistemic Injustice. Power and the Ethics of Knowing.* Oxford: Oxford University Press.

Fuš, M. Manuscript a. Towards Antirealism about Genericity.

—— Manuscript b. Arbitrary Reference Account of Generics.

Gelman, S.A. 2003. *The Essential Child: Origins of Essentialism in Everyday Thought*. New York: Oxford University Press.

—— 2010. Generics as a Window onto Young Children's Concepts. In *Kinds, Things, and Stuff: The Cognitive Side of Generics and Mass Terms*, ed. F.J. Pelletier. New York: Oxford University Press.

Gelman, S.A. and P. Bloom. 2007. Developmental Changes in the Understanding of Generics. *Cognition*, 105(1): 166–183.

Gelman, S.A., Goetz, P.J., Sarnecka, B.S. and J. Flukes. 2008. Generic Language in Parent-Child Conversations. *Language Learning and Development*, 4(1): 1–31.

Gelman, S.A. and L. Raman. 2003. Preschool Children Use Linguistic Form Class and Pragmatic Cues to Interpret Generics. *Child Development*, 74(1): 308–325.

Gendler, T. 2008. Alief and Belief. *Journal of Philosophy*, 105: 634–663.

—— 2011. On the Epistemic Costs of Implicit Bias. *Philosophical Studies*, 156: 33–63.

Geurts, B. 1985. Generics. *Journal of Semantics*, 4.

Gibbard, A. 1990. *Wise Choices, Apt Feelings*. Cambridge: Harvard University Press.

Gooding-Williams, R. 1998. Race, Multiculturalism and Democracy. *Constellations*, 5(1): 18–41.

Goodman, N. 1954. *Fact, Fiction and Forecast*. University of London: Athlone Press.

Graham, S.A., Nayer, S.L. and S.A. Gelman. 2011. Two-Year-Olds Use the Generic/Non-generic Distinction to Guide Their Inferences About Novel Kinds. *Child Development*, 82(2): 493–507.

Green, M. 2020. Speech Acts. In *The Stanford Encyclopedia of Philosophy* (Winter Edition), ed. E.N. Zalta. URL = <https://plato.stanford.edu/archives/win2020/entries/speech-acts/>.

Greene, J. 2008. The Secret Joke of Kant's Soul. In *Moral psychology*, ed. W. Sinnott-Armstrong. Cambridge, MA: MIT Press.

Greenough, P. 2017. Conceptual Marxism and Truth: Inquiry Symposium on Kevin Scharp's Replacing Truth. *Inquiry*, 62(4): 403–421.

—— 2020. Neutralism and Conceptual Engineering. In *Conceptual Engineering and Conceptual Ethics*, eds. Burgess A., Cappelen H. and D. Plunkett. Oxford: Oxford University Press.

—— Manuscript. *Against Conceptual Engineering*.

Grice, H.P. 1957. Meaning. *The Philosophical Review*, 66(3): 377–88.

—— 1975. Method in Philosophical Psychology. *Proceedings and Addresses of the American Philosophical Association*, 68: 23–53.

Halpin, H. and A. Monnin. 2014. *Philosophical Engineering: Toward a Philosophy of the Web*. Wiley Blackwell.

Hardimon, M.O. 2003. The Ordinary Concept of Race. *Journal of Philosophy,* 100(9): 437–55.

Hare, R.M. 1952. *The Language of Morals.* Oxford: Oxford University Press.

Haslanger, S. 1999. What Knowledge Is and What It Ought to Be: Feminist Values and Normative Epistemology. *Noûs,* 33: 459–480.

—— 2000. Gender and Race: (What) Are They? (What) Do We Want Them to Be? *Noûs,* 34(1): 31–55.

—— 2005. What Are We Talking About? The Semantics and Politics of Social Kinds. *Hypatia,* 20(4): 10–26.

—— 2006. What Good Are Our Intuitions? Philosophical Analysis and Social kinds. *Aristotelian Society Supplementary,* 80(1): 89–118.

—— 2011. Ideology, Generics, and Common Ground. In *Feminist Metaphysics,* ed. C. Witt. Springer Verlag.

—— 2012. *Resisting Reality: Social Construction and Social Critique*. Oxford: Oxford University Press.

Heim, I. 1982. The Semantics of Definite and Indefinite Noun Phrases. Ph.D. dissertation, University of Massachusetts, Amherst.

Heyd, D. 2019. Supererogation. In *The Stanford Encyclopedia of Philosophy* (Winter Edition), ed. E.N. Zalta. URL = <https://plato.stanford.edu/archives/win2019/entries/supererogation/>.

Hirsch, E. 2005. Physical-Object Ontology, Verbal Disputes, and Common Sense. *Philosophy and Phenomenological Research*, 70: 67– 98.

—— 2008. Ontological Arguments: Interpretive Charity and Quantifier Variance. In *Contemporary Debates in Metaphysics*, eds. Hawthorne, J., Sider, T. and D. Zimmerman. Cambridge, Mass.: Blackwell.

—— 2009. Ontology and Alternative Languages. In *Meta-metaphysics: New Essays on the Foundations of Ontology*, eds. Chalmers, D., Manley, D. and R. Wasserman. New York: Oxford University Press.

Hollander, M.A., Gelman, S.A. and J. Star. 2002. Children's Interpretation of Generic Noun Phrases. *Developmental Psychology*, 38(6): 883–894.

Horn, L.R. 1989. *A Natural History of Negation*. Chicago: University of Chicago Press.

Horn, L.R. and H. Wansing. 2020. Negation. In *The Stanford Encyclopedia of Philosophy* (Spring Edition), ed. E.N. Zalta. URL = <https://plato.stanford.edu/archives/spr2020/entries/negation/>.

Hornsby, J. and R. Langton. 1998. Free Speech and Illocution. *Journal of Legal Theory,* 4: 21–37.

Horvath, J. 2017. Philosophical Analysis: The Concept Grounding View. *Philosophy and Phenomenological Research*, doi: 10.1111/phpr.12378.

Horvath, J. and S. Koch. 2020. Experimental Philosophy and the Method of Cases. *Philosophy Compass*, doi:/10.1111/phc3.12716.

Huemer, M. 2013. *The Problem of Political Authority: An Examination of the Right to Coerce and the Duty to Obey*. Basingstoke: Palgrave Macmillan.

Ichikawa, J. 2020. Contextual Injustice. *Kennedy Institute of Ethics Journal,* 30 (1): 1–30.

IFLScience. 2020. *Why The WHO Is Now Using The Phrase "Physical Distancing" Instead Of "Social Distancing"* [Internet]. Available from: https://www.iflscience.com/health-and-medicine/why-the-who-is-now-using-the-phrase-physical-distancing-instead-of-social-distancing [Accessed 11 January 2021].

Independent. 2019. *PETA just compared using animal idioms to racism and homophobia and the backlash is real* [Internet]. Available from: https://www.indy100.com/discover/peta-animal-rights-idioms-racism-homophobia-twitter-8668046 [Accessed 11 January 2021].

Isaac, M.G. 2020. How to Conceptually Engineer Conceptual Engineering? *Inquiry*, doi: 10.1080/0020174X.2020.1719881.

—— Manuscript. The Bootstrapping Challenge: What Should Conceptual Engineering Be all About?

Jackson, F. 1998. *From Metaphysics to Ethics: A Defence of Conceptual Analysis*. Oxford: Oxford University Press.

—— 2004. What are Proper Names For? Experience and Analysis. In *Kirchberg Proceedings*, eds. Reicher, M.E. and J.C. Markek. Vienna: Kirchberg.

Jenkins, C. 2014. Merely Verbal Disputes. *Erkenntnis,* 79 (1): 11–30.

Jeshion, R. 2013. Expressivism and the Offensiveness of Slurs. *Philosophical Perspectives*, 27(1): 231–259.

Jorem, S. 2020. Conceptual Engineering and the Implementation Problem. *Inquiry,* doi: 10.1080/0020174X.2020.1809514.

Joyce, R. 2005. *The Myth of Morality.* Cambridge University Press.

—— 2006. *The evolution of morality.* Cambridge, MA: MIT Press.

—— 2013. Ethics and Evolution. In *The Blackwell Guide to Ethical Theory (2nd ed.)*, eds. LaFollette, H. and I. Persson. Malden, MA: Blackwell.

Kearns, S. and O. Magidor. 2012. Semantic Sovereignty. *Philosophy and Phenomenological Research,* 58: 322–350.

Khemlani, S., Leslie, S.J. and S. Glucksberg. 2012. Inferences About Members of Kinds: The Generics Hypothesis. *Language and Cognitive Processes*, 27(6): 887–900.

King, J.C. 1998. What Is a Philosophical Analysis? *Philosophical Studies,* 90: 155–179.

Kitcher, P. 1999. Race, Ethnicity, Biology, Culture. In *Racism*, ed. L. Harris. New York, Humanity Books.

Kitsik, E. 2018. Explication as a Strategy for Revisionary Philosophy. *Synthese*, doi: 10.1007/s11229-018-1774-z.

Koch, S. 2018. The Externalist Challenge to Conceptual Engineering. *Synthese*, doi.org/10.1007/s11229-018-02007-6.

—— 2019. Carnapian Explications, Experimental Philosophy, and Fruitful Concepts. *Inquiry*, 62 (6): 700–717

—— 2020a. Engineering What? On Concepts in Conceptual Engineering. *Synthese*, doi.org/10.1007/s11229-020-02868-w.

—— 2020b. There is no Dilemma for Conceptual Engineering. Reply to Max Deutsch. *Philosophical Studies,* doi.org/10.1007/s11098-020-01546-4.

Korman, D.Z. 2014. Debunking Perceptual Beliefs about Ordinary Objects. *Philosophers' Imprint*, 14(13): 1–21.

Kornblith, H. 2002. *Knowledge and Its Place in Nature.* Oxford: Oxford University Press.

Krifka, M., partly in collaboration with Claudia Gerstner. 1987. *An Outline of Genericity*, *SNS-Bericht 87–23*. University of Tubingen.

Krifka, M., Pelletier, F., Carlson, G., ter Meulen, A., Chierchia, G. and G. Link. 1995. Genericity: An Introduction. In *The Generic Book*, eds. Carlson G.N. and F. J. Pelletier. Chicago: University of Chicago Press.

Kripke, S.A. 1971 Identity and Necessity. In *Identity and Individuation*, ed. M.K. Munitz. New York: New York University Press.

—— 1980. *Naming and Necessity.* Harvard University Press.

Königs, P. 2018. Two Types of Debunking Arguments. *Philosophical Psychology,* 31(3): 383–402.

Lakoff, G. 2004. *Don't Think of an Elephant: Know Your Values and Frame the Debate*. Chelsea Green Publishing.

Langton. R. 1993. Speech Acts and Unspeakable Acts. *Philosophy and Public Affairs*, 22(4): 293–330.

Langton, R., Haslanger, S. and L. Anderson. 2012. Language and Race. In *The Routledge Companion to Philosophy of Language*, eds. Russell, G. and D. Graff Fara. NewYork: Routledge.

Lawler, J. 1972. Generic to a Fault. In *Papers from the 8th Regional Meeting, Chicago Linguistic Society* (CLS 8). Chicago: Chicago Linguistic Society.

Lawler, J. 1973. Studies in English Generics. *University of Michigan Papers in Linguistics 1(1).*

Leslie, S. J. 2007. Generics and the Structure of the Mind. *Philosophical Perspectives*, 375–405.

—— 2008. Generics: Cognition and Acquisition. *Philosophical Review,* 117(1): 1–49.

—— 2012. Generics Articulate Default Generalizations. *Recherches Linguistiques de Vincennes,* 41: 25–44.

—— 2013. Essence and Natural Kinds: When Science Meets Preschooler Intuition. In *Oxford Studies in Epistemology*, vol. 4, eds. T. Gendler and J. Hawthorne. Oxford: Oxford University Press.

—— 2014. Carving up the Social World with Generics. *Oxford Studies in Experimental Philosophy,* Volume 1, eds. Knobe, J., Lombrozo, T. and S. Nichols. Oxford: Oxford University Press.

—— 2017. The Original Sin of Cognition: Fear, Prejudice, and Generalization. *Journal of Philosophy,* 8: 393–421.

Leslie, S.J. and S.A. Gelman. 2012. Quantified Statements are Recalled as Generics: Evidence from Preschool Children and Adults. *Cognitive Psychology*, 64(3): 186–214.

Leslie, S.J. and A. Lerner. 2016. Generic Generalizations. In *The Stanford Encyclopedia of Philosophy* (Winter Edition), ed. E.N. Zalta. URL = <https://plato.stanford.edu/archives/win2016/entries/generics/>.

Lewis, D. 1975. Adverbs of Quantification. In *Formal Semantics of Natural Language*, ed. E. Keenan. Cambridge: Cambridge University Press.

—— 1984. Putnam's Paradox. *Australasian Journal of Philosophy*, 62: 221–236.

Liebesman, D. 2011. Simple Generics. *Noûs*, 45(3): 409–442.

Ludlow, P. 2014. *Living Words: Meaning Underdetermination and the Dynamic Lexicon*. Oxford: Oxford University Press.

Lycan, W.G. 1999. *Philosophy of Language: A Contemporary Introduction*. Routledge.

Löhr, G. 2018. Concepts and Categorization: Do Philosophers and Psychologists Theorize About Different Things? *Synthese*, doi:10.1007/s11229-018-1798-4.

Machery, É. 2009. *Doing Without Concepts.* Oxford: Oxford University Press.

—— 2017. *Philosophy within its Proper Bounds.* Oxford: Oxford University Press.

MacFarlane, J. 2011. What Is Assertion? In *Assertion: New Philosophical Essays*, eds. Brown, J. and H. Cappelen. Oxford: Oxford University Press.

MacKinnon, C. 1992. Pornography, Civil Rights and Speech. In *Pornography: Women, Violence and Civil Liberties*, ed. C. Itzin. Oxford: Oxford University Press.

Madva, A. 2016. Virtue, Social Knowledge, and Implicit Bias. In *Implicit Bias & Philosophy:* Volume 1, *Metaphysics and Epistemology*, eds. Brownstein, M. and J. Saul. Oxford: Oxford University Press.

Mallon, R. 2004. Passing, Traveling and Reality: Social Constructionism and the Metaphysics of Race. *Noûs*, 38(4): 644–673.

Manley, D. 2009. Introduction. In *Metametaphysics: New Essays on the Foundations of Ontology*, eds. Chalmers, D., Manley, D. and R. Wasserman. New York: Oxford University Press.

Manne, K. 2017. *Down Girl: The Logic of Misogyny.* Oxford: Oxford University Press.

Margolis, E, and S. Laurence. 2010. Concepts and Theoretical Unification. *Behavioral and Brain Sciences*, 33(2&3): 219–220.

Mason, K. 2010. Debunking Arguments and the Genealogy of Religion and Morality. *Philosophy Compass*, 5(9): 770–778.

McCarthy, J. 1986. Applications of Circumscription to Formalizing Commonsense Knowledge. *Artificial Intelligence*, 28: 89–116.

McDowell, J. 2009. *Having the World in View.* Cambridge, MA: Harvard University Press.

McGowan, M.K. 2009. Oppressive Speech. *Australasian Journal of Philosophy*, 87(3): 389–407.

McPherson, T. and D. Plunkett. 2020. Conceptual Ethics and the Methodology of Normative Inquiry. In *Conceptual Engineering and Conceptual Ethics*, eds. Burgess A., Cappelen H. and D. Plunkett.

Oxford: Oxford University Press.

Moore, G.E. 1922. *Philosophical Studies*. London: Routledge.

Morton, J. 2016. A New Evolutionary Debunking Argument Against Moral Realism. *Journal of the American Philosophical Association*, 2(2): 233–253.

Moskowitz, G. and P. Li. 2011. Egalitarian Goals Trigger Stereotype Inhibition: A Proactive Form of Stereotype Control. *Journal of Experimental Social Psychology*, 47(1): 103–116.

Moty, K. and M. Rhodes. Forthcoming. The Unintended Consequences of the Things We Say: What Generics Communicate to Children about Unmentioned Categories. *Psychological Science.*

Murphy, G. 2004. *The Big Book of Concepts*. Cambridge, MA: MIT Press.

Nado, J. 2019. Conceptual Engineering, Truth, and Efficacy. S*ynthese*, doi.org/10.1007/s11229-019-02096-x.

—— 2020: Taking Control: Conceptual Engineering Without (much) Metasemantics. *Inquiry*, doi.org/10.1080/0020174X.2020.1850342.

National Aeronautics and Space Administration. 2019. *2.0 Fundamentals of Systems Engineering* [Internet]. Available from: https://www.nasa.gov/seh/2-fundamentals [Accessed 11 January 2021].

Neale, S. 2000. On Being Explicit: Comments on Stanley and Szábo, and on Bach. *Mind and Language*, 15(2&3): 284–94.

Nickel, B. 2008. Generics and the Ways of Normality. *Linguistics and Philosophy*, 31(6): 629–48.

—— 2010. Ceteris Paribus Laws: Generics and Natural Kinds. *Philosopher's Imprint*, 10(6): 1–25.

—— 2016. *Between Logic and the World: An Integrated Theory of Generics*. Oxford: Oxford University Press.

—— 2017. Generics. In *A Companion to the Philosophy of Language*. 2nd edition, eds. Hale, B., Wright, C. and A. Miller. Malden, Mass.: Wiley Blackwell.

Nozick, R. 1981. *Philosophical Explanations.* Cambridge, MA: Harvard University Press.

Ojeda, A. 1991. Definite Descriptions and Definite Generics. *Linguistics and Philosophy*, 14: 367–397.

Omi, M. and H. Winant. 1994. Racial Formation. In *Racial Formation in the United States*, eds. Omi M. and H. Winant. New York: Routledge.

Papineau, D. 2020. Naturalism. In *The Stanford Encyclopedia of Philosophy*, ed. E.N. Zalta. <https://plato.stanford.edu/archives/sum2020/entries/naturalism/>.

Peacocke, C. 1992. *A Study of Concepts*. Cambridge, MA: MIT Press.

Pelletier, F.J. and N. Asher. 1997. Generics and Defaults. In *Handbook of Logic and Language*, eds. van Benthem, J. and A. ter Meulen. Cambridge, MA: MIT Press.

Pinder, M. (2019). Conceptual Engineering, Metasemantic Externalism and Speaker-Meaning. *Mind*, doi: 10.1093/mind/fzz069.

––– 2020. Conceptual Engineering, Speaker-meaning and Philosophy. *Inquiry*, doi.org/10.1080/0020174X.2020.1853342.

—— Manuscript. Phenomenal Conceptual Engineering.

Pitt, D. 1999. In Defense of Definitions. *Philosophical Psychology,* 12(2): 139–156.

Plunkett, D. 2015. Which Concepts Should We Use?: Metalinguistic Negotiations and The Methodology of Philosophy. *Inquiry,* 58(7&8): 828–874.

—— 2016. Conceptual History, Conceptual Ethics, and The Aims of Inquiry: A Framework for Thinking About The Relevance of the History/Genealogy of Concepts to Normative Inquiry. *Ergo*, 3(2): 27–64.

Plunkett, D. and T. Sundell. 2013. Disagreement and the Semantics of Normative and Evaluative Terms. *Philosophers' Imprint*, 13(23): 1–37.

Pollock, J. 2020. Content Internalism and Conceptual Engineering. *Synthese*, doi.org/10.1007/s11229-020-02815-9.

Prasada, S. and E.M. Dillingham. 2006. Principled and Statistical Connections in Common Sense Conception. *Cognition*, 99: 73–112.

Prasada, S., Sangeet K., Leslie S.J. and S. Glucksberg. 2013. Conceptual distinctions amongst generics. *Cognition,* 126: 405–422.

—— 2009. Representation of Principled Connections: A Window onto the Formal Aspect of Common Sense Conception. *Cognitive Science,* 33(3): 401–448.

Prinzing, M. 2018. The Revisionist's Rubric: Conceptual Engineering and the Discontinuity Objection. *Inquiry*, 61(8): 854–880.

Quine, W.V.O. 1951. Two Dogmas of Empiricism. *Philosophical Review,* 60(1): 20–43.

—— 1960. *Word and Object*. The MIT Press. Cambridge Massachusetts Railton.

Railton, P. 1989. Naturalism and Prescriptivity. *Social Philosophy and Policy,* 7(1): 151–174.

—— 1993. Noncognitivism about Rationality: Benefits, Costs, and an Alternative. *Philosophical Issues,* 4: 36–51.

Recanati, F. 1989. The Pragmatics of What is Said. *Mind and Language,* 4(4): 295–329.

—— 2004. *Literal Meaning*. Cambridge, Mass.: Cambridge University Press.

Reiter, R. 1987. Nonmonotonic Reasoning. In *Annual Review of Computer Science*, eds. Traub, J.F., Nilsson, N.J. and B.J. Grosz. Palo Alto: Annual Reviews Inc.

Rey, G. 1985. Concepts and Conceptions: A Reply to Smith, Medin and Rips. *Cognition,* 19(3): 297–303.

—— 2010. Concepts Versus Conceptions (Again). *Behavioral and Brain Sciences,* 33(2&3): 221–222.

Rhodes, M., Leslie, S.J. and C. Tworek. 2012. Cultural Transmission of Social Essentialism. *Proceedings of the National Academy of Sciences (PNAS)*, 109(34): 13526–13531.

Riggs, J. 2019. Conceptual Engineers Shouldn't Worry about Semantic Externalism. *Inquiry*, online-first. doi.org/10.1080/0020174X.2019.1675534.

Ritchie, K. 2019. Should We Use Racial and Gender Generics? *Thought: A Journal of Philosophy,* 8: 33–41.

Saul, J. 2017. Are Generics Especially Pernicious? *Inquiry,* doi:10.1080/0020174X.2017.1285995.

—— 2018. Dogwhistles, Political Manipulation, and the Philosophy of Language. In *New Work on Speech Acts*, eds. Daniel F., Matt C. and D. Harris. Oxford: Oxford University Press.

Sawyer, S. 2018. The Importance of Concepts. *Proceedings of the Aristotelian Society*, 116(2): 127–147.

—— 2020a. Talk and Thought. In *Conceptual Engineering and Conceptual Ethics*, eds. Burgess A., Cappelen H. and D. Plunkett. Oxford: Oxford University Press.

—— 2020b. The Role of Concepts in Fixing Language. *Canadian Journal of Philosophy,* 50(5): 555–565.

Scharp, K. 2007. Replacing truth. *Inquiry,* 50(6): 606–21.

—— 2013. *Replacing Truth.* Oxford: Oxford University Press.

—— 2020. Philosophy as the Study of Defective Concepts. In *Conceptual Engineering and Conceptual Ethics*, eds. Burgess A., Cappelen H. and D. Plunkett. Oxford: Oxford University Press.

Scharp, K. and S. Shapiro. 2017. Revising Inconsistent Concepts. In *Reflections on the Liar*, ed. B. Armour-Garb. Oxford: Oxford University Press.

Schubert, L.K. and F.J. Pelletier. 1987. Problems in Representing the Logical Form of Generics, Bare Plurals, and Mass Terms. In *New Directions in Semantics,* ed. E. Lenore. Academic Press.

—— 1989. Generically Speaking, or Using Discourse Representation Theory to Interpret Generics. In *Properties, Types and Meaning. Semantic Issues,Vol. II,* eds. Chierchia, G., Partee, B.H. and R. Turner. Dordrecht: Kluwer.

Searle, J. 1969. *Speech Acts: An Essay in the Philosophy of Language.* Cambridge, Cambridge University Press.

—— 1995. *The Construction of Social Reality.* Free Press.

Shadbolt N. 2007. Philosophical Engineering. In *Words and Intelligence II. Text, Speech and Language Technology,* vol 36., eds. Ahmad K., Brewster C. and M. Stevenson. Springer, Dordrecht.

Shafer-Landau, R. 2003. *Moral Realism: A Defence.* Oxford: Oxford University Press.

Shepard, J. and J. Justus. 2015. X-Phi and Carnapian explication. *Erkenntnis,* 80: 381–402.

Shields, M. (2020): Conceptual Change in Perspective. *Inquiry*, doi: 10.1080/0020174X.2020.1805705.

Simion, M. 2018. The 'Should' in Conceptual Engineering. *Inquiry*, 61(8): 914–928.

Simion, M. and C. Kelp. 2020. Conceptual Innovation, Function First. *Nous,* 54(4): 985–1002.

Soames, S. 2003. *Philosophical Analysis in the Twentieth Century*, Volume 1: *The Dawn of Analysis*, Volume 2: *The Age of Meaning*. Princeton University Press.

—— 2005. Philosophical Analysis. In *Encyclopedia of Philosophy*, ed. D.M. Borchert, 2nd ed. Detroit: Thomson Gale.

Stanley, J. and Z. Szábo. 2000. On Quantifier Domain Restriction. *Mind and Language,* 15, (2&3): 219–61.

Sterelny, K. 2012. *The Evolved Apprentice: How Evolution Made Humans Unique.* Cambridge, MA, MIT Press.

Sterken, R. 2015. Generics in Context. *Philosopher's Imprint*, 15(21): 1–30.

—— 2016. Between Logic and the World: An Integrated Theory of Generics, by Bernhard Nickel. *Notre Dame Philosophical Reviews. University of Notre Dame* [Internet]. Available from: https://ndpr.nd.edu/news/between-logic-and-the-world-an-integrated-theory-of-generics [Accessed 11 January 2021].

—— 2020. Linguistic Interventions and Transformative Communicative Disruption. In *Conceptual*

Engineering and Conceptual Ethics, eds. Burgess A., Cappelen H. and D. Plunkett. Oxford: Oxford University Press.

Strawson, P.F. 1963. Carnap's Views on Constructed Systems Versus Natural Languages in Analytic Philosophy. In *The Philosophy of Rudolf Carnap*, ed. P. Schilpp. Lasalle, IL: Open Court.

—— 1964. Intention and Convention in Speech Acts. *The Philosophical Review*, 73(4): 439–460.

Street, S. 2006. A Darwinian Dilemma for Realist Theories of Value. *Philosophical Studies, 127*(1): 109–166.

Tanswell, F. 2017. Conceptual Engineering for Mathematical Concepts. *Inquiry*, doi: 10.1080/0020174X.2017.1385526.

Tardif, T., Gelman, S.A., Fu, X. and L. Zhu. 2011. Acquisition of Generic Noun Phrases in Chinese: Learning About Lions Without an '-s'. *Journal of Child Language*, 30(1): 130–161.

Tarski, A. 1933. The Concept of Truth in Formalized Languages. In *Logic, Semantics, Meta-Mathematics*, ed. J. Corcoran. Indianapolis: Hackett Publishing Company, 1983.

Thakral, R. Manuscript. Moral Principles.

Thomasson, A. L. 2015. *Ontology Made Easy.* New York: Oxford University Press.

—— 2016. Metaphysical Disputes and Metalinguistic Negotiation. *Analytic Philosophy*, 57(3): 1–28.

—— 2020. A Pragmatic Method for Conceptual Ethics. In *Conceptual Engineering and Conceptual Ethics*, eds. Burgess A., Cappelen H. and D. Plunkett. Oxford: Oxford University Press.

Thurow, J.C. 2013. Does Cognitive Science Show Belief in God to Be Irrational? The Epistemic Consquences of the Cognitive Science of Religion. *International Journal for Philosophy of Religion*, 74(1): 77–98.

Tye, M. 1992. Naturalism and the Mental. *Mind,* 101(403): 421–441.

van der Kolk, B.A. 2014. *The body keeps the score: Brain, mind, and body in the healing of trauma.* Viking.

van der Schaar, M. 2013. Judgement, Propositional Attitudes and the Proposition (1908–1944). In *G.F. Stout and the Psychological Origins of Analytic Philosophy*. Palgrave Macmillan, London.

Walker, M. 2017. *Why We Sleep: Unlocking the Power of Sleeping and Dreams*. Scribner, an imprint of Simon & Schuster.

Weatherson, B. 2003. What Good Are Counterexamples? *Philosophical Studies,* 115(1): 1–31.

Weinberg, J.M. 2006. What's Epistemology For? The Case for Neopragmatism in Normative

Metaepistemology. In *Epistemological Futures*, ed. S. Hetherington. Oxford: Clarendon Press.

Westerståhl, D. Generalized Quantifiers. 2016. In *The Stanford Encyclopedia of Philosophy* (Winter Edition), ed. E.N. Zalta. URL = <https://plato.stanford.edu/archives/win2016/entries/generalized-quantifiers/>.

Williamson T. 2007. *The Philosophy of Philosophy.* Blackwell Pub.

—— 2009. Reference, Inference, and the Semantics of Pejoratives. In *The Philosophy of David Kaplan*, eds. Almog, J. and P. Leonardi. Oxford: Oxford University Press.

Wodak, D. and S.J. Leslie. 2017. The Mark of the Plural: Generic Generalizations and Race. In *The Routledge Companion to the Philosophy of Race*, eds. Taylor, P.C., Alcoff, L.M. and L. Anderson. New York: Routledge.

Wodak, D., Leslie, S. J., and M. Rhodes. 2015. What a Loaded Generalization: Generics and Social Cognition. *Philosophy Compass,* 10(9): 625–634.

Woodward, J. 2003. *Making Things Happen: A Theory of Causal Explanation*. Oxford: Oxford University Press.

Young, J.W. 1991. *Totalitarian Language: Orwell's Newspeak and its Nazi and Communist Antecedents*. Charlottesville: University Press of Virginia.

Zack, N. 1997. Race and Philosophic Meaning. In *RACE/SEX: Their Sameness, Difference, and Interplay*. New York: Routledge.

—— 2002. *Philosophy of Science and Race.* New York: Routledge.

SUMMARY (ENGLISH)

Assert This: "Philosophers Are Engineers"

A Study of Philosophical Engineering and Generic Judgments

Can the pernicious effects of generic statements such as "Blacks are violent," "Women are submissive," or "Muslims are terrorists" be ameliorated by applying the method of engineering in philosophy? This dissertation brings together and contributes to recent debates on *conceptual engineering* and *social group generics*.

Part I of this dissertation does some groundwork for the philosophical method of engineering in philosophy, its *nature* and *terminology*. It makes a contribution to the field of philosophical methodology by offering a general methodological account of engineering in philosophy I dub '*philosophical engineering*.' In particular, I motivate and develop my account by arguing for a *five stage* (identification, evaluation, strategic planning, implementation, re-evaluation) *recursive model* of philosophical engineering. I take this to be congenial with other, non-philosophical branches of engineering. Furthermore, my account endorses *pluralism about the objects of engineering in philosophy*, allowing, thus, for a possibility of engineering different philosophical objects, such as terms, concepts, speaker-meanings, topics, speech acts, theses, arguments, properties, etc.

Part II of this dissertation brings to the foreground *generic judgments*, their relation to *generic truths* and their role in *ameliorating* the morally, socially and politically pernicious effects of generics. It makes a contribution both to the field of generics and the field of philosophical methodology by developing a domain-specific account for *engineering generic judgments*. First, I offer a *debunking argument for realist theories of genericity* by arguing that an alternative, evolutionary explanation of the role that our generic judgments play in connection to the independent generic truths beats the realist's account on scientific merits. Second, I suggest an ameliorative engineering project for the pernicious effects of generics which takes generic judgments as the main object of engineering and calls for correction of generic judgments. As one of the plausible strategies to achieve engineering of generic judgments, I suggest the *Speech Act Approach*, an approach to changing doxastic propositional attitudes towards generics by utilizing speech acts. The proposed approach is independent from any particular view about the semantics, epistemology and metaphysics of generics.

SAMMENDRAG (NORSK)

Påstå Dette: «Filosofer Er Ingeniører»

En Studie av Filosofisk Ingeniørvirksomhet og Generiske Dommer

Kan vi oppheve de skadevirkningene til generiske utsagn av typen «svarte er voldelige», «kvinner er underdanige», eller «muslimer er terrorister», gjennom filosofisk ingeniørkunst? Denne avhandlingen bidrar til og sammenstiller nylige debatter om henholdsvis *begrepsrettet ingeniørvirksomhet* (eng. conceptual engineering) og *generiske uttrykksformer om sosiale grupper*.

Avhandlingens første del bidrar til det grunnarbeidet som foregår med å avklare filosofisk ingeniørkunst, både hva en slik *metode* består i og hva slags *terminologi* den best artikuleres med. Denne delen bidrar til den filosofiske metode, ved å gi en generell metodologisk teori om ingeniørkunst i filosofi, noe jeg kaller «*filosofisk ingeniørkunst*» (eng. 'philosophical engineering'). Beskrivelsen tjener som en generell metodologisk redegjørelse for revisjonisme i filosofi. Jeg argumenterer for en *fem-trinns rekursiv modell* som innebefatter *identifikasjon*, *evaluering*, *strategisk planlegging*, *implementering* og *re-evaluering*. Stegene i modellen har klare paralleller med trekk i ordinær, ikke-filosofisk ingeniørvirksomhet. Jeg forsvarer et *pluralistisk syn om hva slags størrelser som kan være gjenstand for filosofisk ingeniørvirksomhet.* Jeg argumenterer altså for at det er forskjellige typer ting man kan ta sikte på å revidere eller konstruere som filosof; slik som språklige termer, begreper, talers mening (eng. speaker-meaning), tema, talehandlinger, teser, argumenter, egenskaper osv.

Avhandlingens andre del tar for seg *generiske dommer*, deres forbindelse til *generiske sannheter* og hvilken rolle de har i *å lege eller oppheve* de moralske, sosiale og politiske skadevirkningene til generiske utsagn. Jeg gir en domenespesifikk redegjørelse for hvordan vi kan *revidere generiske dommer*, og bidrar således både til litteraturen om generiske uttrykksformer og til filosofisk metodologi. Først gir jeg et *argument som undergraver realistiske teorier om generisitet* ved å vise at det finnes en evolusjonær forklaring på forbindelsen mellom generiske dommer og generiske sannheter som overgår realistens forklaring. Dernest foreslår jeg at vi angriper de skadelige virkningene til generiske utsagn med et revisjonsprosjekt som tar generiske dommer som den spesifikke gjenstanden for revisjon. Det er flere plausible måter å revidere generiske dommer på. Jeg foreslår *talehandlingsmetoden*, en metode som går ut på å bruke talehandlinger til å endre våre trosoppfatninger om påstandsinnholdet som blir fremmet med et generisk utsagn. Dette prosjektet er kompatibelt med ulike syn på de metafysiske, erkjennelsesteoretiske og semantiske egenskapene til generiske uttrykksformer.

"Would you tell me, please, which way I ought to go from here?"

"That depends a good deal on where you want to get to," said the Cat.

"I don't much care where—" said Alice.

"Then it doesn't matter which way you go," said the Cat.

"—so long as I get SOMEWHERE," Alice added as an explanation.

"Oh, you're sure to do that," said the Cat, "if you only walk long enough."

~ *Alice in Wonderland*, Lewis Carroll